# Hybrid GNN-Centric Architectures for AI-Native 6G Wireless Networks: A Comprehensive Survey

Mostafa Rahmani Ghourtani, Sajedeh Norouzi, Jinxuan Chen, Hamed Ahmadi, Torsten Braun, *Senior Member, IEEE*, Kaushik Chowdhury, *Fellow, IEEE*, and Alister Burr, *Life Senior Member, IEEE*

***Abstract*—The growing complexity, scale, and heterogeneity of 6G wireless systems call for a shift toward AI-native architectures that are not only data-driven but also topology-aware, adaptive, and distributed. Graph Neural Networks (GNNs), with their native support for graph-structured data, are well-suited for modeling the irregular and dynamic relationships inherent in wireless communication systems. However, standalone GNNs may be insufficient to address key 6G challenges such as continual learning, data scarcity, and dynamic adaptation. This survey, therefore, explores the emerging synergy between GNNs and complementary AI paradigms, including deep reinforcement learning (DRL), federated learning (FL), meta-learning, generative models, mixture-of-experts (MoE), and world models, enabling hybrid AI–GNN architectures for intelligent control, predictive adaptation, and scalable optimization across the wireless stack. We systematically review how these GNN-centric AI models can support intelligent functionality in next-generation network architectures such as O-RAN, as well as core 6G domains, including edge computing for wireless systems, advanced MIMO, traffic prediction, and digital twins. The survey also highlights key challenges in hybrid AI–GNN adoption, particularly scalability, generalization across dynamic topologies, interpretability, and symbolic reasoning, and discusses emerging strategies such as graph causality learning, dynamic GNNs, and neurosymbolic integration to address them. By consolidating recent advances and outlining open research directions, this work positions hybrid AI–GNN architectures as a promising approach toward enabling intelligent, energy-efficient, and context-aware wireless systems for 6G and beyond.**

***Index Terms*—Hybrid GNN, distributed and heterogeneous systems, AI-native networks, wireless networks.**

## I. Introduction

SIXTH Generation (6G) wireless networks are expected to extend beyond improvements in speed and latency, enabling communication architectures that support immersive and intelligent applications such as holographic telepresence and extended reality (XR) [1], [2]. Meeting the anticipated complexity and scale of 6G systems will require networks that can learn, adapt, and optimize with minimal human intervention.

In this context, Artificial Intelligence (AI) is expected to play a central role in enabling data-driven and adaptive network operation. It is important to distinguish between different levels of AI integration in wireless systems. *AI-enabled* approaches employ AI as an auxiliary tool to enhance specific functions, while the overall system architecture remains largely unchanged. *AI-driven* systems rely more heavily on AI for decision-making, where learning-based models influence control and optimization but are still embedded within predefined system structures. In contrast, *AI-native* architectures fundamentally integrate AI into the network design itself, embedding learning and inference directly into control loops and system operation. This distinction is particularly relevant for 6G, where AI-native design implies tight coupling between data, models, and distributed network intelligence.

AI-native architectures aim to embed intelligence into the network fabric, supporting functionalities such as self-optimization, demand prediction, and real-time service personalization [3].

Graph Neural Networks (GNNs) are particularly well suited for wireless networks due to their ability to model non-Euclidean, graph-structured data, such as interference, mobility, and cooperation graphs [4]. This enables GNNs to capture relational dependencies that are central to tasks such as resource allocation, handover, and distributed caching. Recent studies have demonstrated the effectiveness of GNNs in scenarios where relational structure plays a key role. For example, in resource allocation and power control, GNNs can capture spatial dependencies between users and base stations more effectively than fully connected or convolutional models. In user association and scheduling, GNN-based approaches have shown improved adaptability compared to rule-based methods under dynamic topology and traffic conditions. Furthermore, GNNs can offer scalability advantages in large-scale distributed settings by enabling localized message passing and inference [5].

Building on these properties, recent work has explored integrating GNNs with other learning paradigms to address the heterogeneous demands of 6G systems. These integrations give rise to what we refer to as **Hybrid GNN** architec-

Received 10 January 2026; revised 1 April 2026; accepted 1 April 2026. Date of publication 6 April 2026; date of current version 14 April 2026. This work was supported in part by the Department for Science, Innovation and Technology, U.K., under the Yorkshire Open-RAN (YO-RAN) project under Grant TS/X013758/1; in part by the U.K. Engineering and Physical Sciences Research Council (EPSRC) under the HASC HUB ORLANDO project under Grant EP/X040569/1; in part by the Swiss National Science Foundation under Grant 219330; and in part by the U.S. National Science Foundation under Grant ECCS-2516080 and Grant CNS-2112471. *(Corresponding author: Mostafa Rahmani Ghourtani.)*

Mostafa Rahmani Ghourtani, Hamed Ahmadi, and Alister Burr are with the School of Physics, Engineering and Technology, University of York, YO10 5DD York, U.K. (e-mail: rahmani.mostafa@york.ac.uk; hamed.ahmadi@york.ac.uk; alister.burr@york.ac.uk).

Sajedeh Norouzi, Jinxuan Chen, and Torsten Braun are with the Institute of Computer Science, University of Bern, 3012 Bern, Switzerland (e-mail: sajedeh.norouzi@unibe.ch; jinxuan.chen@unibe.ch; torsten.braun@unibe.ch).

Kaushik Chowdhury is with the Electrical and Computer Engineering Department, The University of Texas at Austin, Austin, TX 78758 USA (e-mail: kaushik@utexas.edu).

Digital Object Identifier 10.1109/COMST.2026.3681198

© 2026 The Authors. This work is licensed under a Creative Commons Attribution 4.0 License.
For more information, see https://creativecommons.org/licenses/by/4.0/

tures, or **Artificial Intelligence–Graph Neural Network (AI–GNN) Synergy**. In this survey, the term **AI–GNN** denotes *GNN-centric architectures integrated with complementary AI paradigms*, rather than implying that GNNs operate independently from AI. Such hybrid models combine GNNs with approaches such as reinforcement learning (RL) for sequential decision-making [6], federated learning (FL) for privacy-preserving distributed learning [7], and explainable AI (XAI) for improved transparency and interpretability [8]. These combinations extend the applicability of GNNs to settings involving temporal dynamics, distributed intelligence, and interpretability, while acknowledging that GNNs alone may not fully address all system requirements.

### A. Existing Reviews and Surveys

As the vision for 6G advances, the integration of AI into wireless networks has emerged as a key research direction. Numerous surveys have explored this intersection, generally falling into three categories: system architecture, algorithmic innovation, and network-level optimization.

*1) Ai-Driven System Architectures for 6G:* Several works emphasize embedding AI into the fabric of 6G infrastructure. Seminal works, such as [9], envision AI-native systems that seamlessly unify terrestrial, aerial, and space-based communication infrastructures, enabling highly autonomous and adaptive network behaviour. Further, [10] delves into the practical challenges of deploying AI at the network edge, with a particular emphasis on achieving a balance between energy efficiency and low-latency inference, which is critical for real-time applications in next-generation wireless systems. Other studies extend these ideas to non-terrestrial networks (NTNs), where AI is envisioned as a key enabler for managing orbital dynamics and satellite coordination [11], [12]. At the physical layer, surveys such as [13] and [14] explore how AI enhances reconfigurable surfaces, joint communication-sensing, and other novel paradigms.

*2) Advanced Learning Models and Ai Techniques:* A second focus in the literature centers on algorithmic innovations, especially scalable and context-aware AI models. Sun et al. [15] highlight the importance of interpretability in AI systems for dynamic resource allocation. Others examine large transformer models for beamforming [16], [17] and meta-learning for rapidly changing environments [18], [19]. Broad surveys such as [20] and [21] present AI applications across the protocol stack, from signal processing to network control.

*3) Network Optimization and Distributed Intelligence:* The third wave of surveys emphasizes AI's role in end-to-end network optimization. Classic optimization is being merged with data-driven intelligence to enhance spectrum efficiency, reduce energy consumption, and improve decision-making under uncertainty [22]. RL, in particular, has gained attention for network slicing [23] and related works examine how AI can enable dynamic slicing, user-centric Radio Access Network (RAN) design, and near-autonomous control under 3GPP standards [14]. Surveys also explore multi-agent reinforcement learning (MARL) [24], federated RL [25], and edge-based learning frameworks [26], particularly for Internet of Things (IoT) systems that demand both privacy and adaptability [27]. Comprehensive surveys, such as [28], highlight AI as a pivotal enabler for building self-optimizing, context-aware wireless ecosystems, emphasizing its central role in driving automation and adaptability across future network architectures.

*4) Graph-Based Learning in Wireless Communication:* In parallel, a separate line of research has focused on graph-theoretic and graph learning methods. Early studies modeled interference and mesh topologies using classical graph theory [29], [30]. Majeed et al. [31] and He et al. [32] expanded this to broader network structures and GNN-based tasks like channel estimation and resource allocation.

Recent surveys provide deeper insights into GNN applications across network functions. Jiang et al. [33] classify GNN models by learning objectives, while Tam et al. [34] discuss their integration with Software Defined Network (SDN)/Network Functions Virtualization (NFV) for programmable core networks. In IoT security and anomaly detection, [35], [36] illustrate GNNs' role in decentralized analytics. Ivanov et al. extend these methods to NTNs [37], and Suarez et al. [38] examine their use in intrusion detection and data fusion. Dai et al. offer perhaps the most comprehensive reviews to date, spanning aerial planning and GNN-powered resource control [39], [40].

More recently, Zeng et al. [41] introduced the concept of *Edge Graph Intelligence (EGI)*, providing a system-level perspective on the reciprocal interplay between graph intelligence and edge networks. Their survey organizes the field along multiple dimensions, including edge applications of graph intelligence, edge infrastructures for training and inference, graph-driven optimization of edge systems, and supporting EGI ecosystems. This work highlights the importance of deployment-aware design and the role of graph intelligence in enabling scalable and adaptive edge-network operations. In addition, Lu et al. [42] present a comprehensive survey on *agentic GNNs*, where multiple GNN models are orchestrated through agent-based intelligence toward edge general intelligence. Their work focuses on graph representation, GNN architectures, and task-level alignment across wireless communication scenarios.

Despite these advances, existing surveys typically adopt either a *GNN-centric* or *edge-/system-centric* perspective, focusing on representation learning, architectural design, or deployment aspects in isolation. Comparatively less attention has been given to a unified view that examines how GNNs can be effectively integrated with complementary AI paradigms. As wireless networks evolve toward more dynamic and distributed operation, understanding such integrations can provide additional insights into designing adaptive and scalable learning frameworks.

### B. Contribution

This survey presents a comprehensive review of GNNs as a promising approach for enabling AI-native 6G wireless networks. Distinct from prior surveys, it emphasizes the synergistic integration of GNNs with emerging AI paradigms, such as FL, XAI, generative models, and others, within

TABLE I
COMPARISON OF EXISTING SURVEYS AND OUR CONTRIBUTIONS

| Survey | Primary Focus | GNN Coverage | AI–GNN Synergy | 6G Relevance | Architectural View | Main Limitation |
|---|---|---|---|---|---|---|
| [33] | GNN models for network management | ✓ | ✗ | ∼ | ✗ | Lacks AI-native 6G integration |
| [34] | GNNs in SDN/NFV | ✓ | ∼ | ✗ | ∼ | Limited to core network, no 6G context |
| [35] | GNNs for IoT security | ✓ | ✗ | ∼ | ✗ | Focuses only on IoT security |
| [36] | GNNs for IoT sensing | ✓ | ✗ | ✗ | ✗ | Narrow sensing scope; no AI-6G synergy |
| [37] | Graph optimization in NTN | ∼ | ✗ | ∼ | ✗ | No GNN/AI integration |
| [5] | Theoretical graph models in wireless | ✓ | ✗ | ✗ | ∼ | Largely theoretical, no practical synergies |
| [38] | GNNs for communication networks | ✓ | ✗ | ∼ | ✗ | Limited to specific use cases |
| [43] | GNNs for IoT resilience | ✓ | ✗ | ∼ | ✗ | No 6G architectural insights |
| [4] | GNNs for mMIMO / aerial networks | ✓ | ∼ | ∼ | ✗ | Technology-specific, no AI-native 6G |
| [39], [40] | Graph-based resource management | ✓ | ✗ | ∼ | ✗ | Graph optimization only; no GNN-AI synergy |
| [6] | GNN-DRL in smart infrastructures | ✓ | ∼ | ∼ | ∼ | Focuses on a single synergy type |
| [44] | GNN + digital twin (DT)s | ✓ | ∼ | ∼ | ∼ | Only DT integration; narrow scope |
| [41] | Edge Graph Intelligence | ✓ | ✗ | ∼ | ✓ | Edge-centric; lacks AI-native 6G integration and no GNN-AI synergy |
| [42] | Agentic GNNs for wireless networks | ✓ | ∼ | ∼ | ✓ | GNN-centric; lacks integration with broader AI paradigms (e.g., DRL, FL) |
| **Ours** | **Hybrid AI–GNN architectures for 6G** | ✓ | ✓ | ✓ | ✓ | **Systematic AI–GNN synergies, 6G enablers, challenges, and future vision** |

*Legend:* ✓ = addressed; ∼ = partially addressed; ✗ = not addressed.

the context of 6G-specific challenges, including scalability, energy efficiency, real-time responsiveness, and heterogeneous architectures. As wireless networks evolve toward greater autonomy, heterogeneity, and real-time adaptability, understanding the role of hybrid GNN architectures at the system level becomes increasingly important.

Table I highlights the distinctions between our work and prior surveys on GNNs in wireless networks. For clarity, we define "AI–GNN Synergy" as the integration of GNNs with other AI paradigms (e.g., deep reinforcement learning (DRL), FL, and generative models), "6G Relevance" as the extent to which works address key characteristics envisioned for 6G networks, including AI-native design, scalability, real-time adaptability, and distributed intelligence across edge–cloud continua, as well as emerging paradigms such as cell-free architectures, Integrated Sensing and Communication (ISAC), and network DTs, and "Architectural View" as the presence of a system-level perspective on integrating GNNs within network architectures.

As shown in Table I, existing surveys typically focus on individual topics such as IoT, sensing, or DTs, and often adopt either a GNN-centric, application-specific, or edge-centric perspective. As a result, they do not fully capture a system-level view of hybrid AI–GNN architectures or their alignment with emerging 6G requirements. In particular, the interaction between GNNs and complementary AI paradigms such as DRL, FL, and neuro-symbolic learning remains only partially explored, and architectural considerations for scalable and adaptive intelligence are often not addressed in a unified manner.

In contrast, this survey provides a unified perspective on AI–GNN integration, covering representative 6G enablers, architectural trends, open challenges, and future research directions.

### C. Paper Organization

The remainder of this survey is structured as follows and is visually summarized in Figure 1. Section II provides a contextual foundation by reviewing the evolution from earlier wireless generations to Fifth Generation (5G), and the anticipated transition to 6G, including a comparative analysis and discussion of emerging challenges. Section III offers an in-depth exploration of GNNs, covering their theoretical foundations, architectural components, model evolution, and recent advances in graph learning, with an emphasis on their alignment with the structure and dynamics of wireless systems. The practical applications of GNNs in wireless communication, spanning resource management, physical layer optimization, security, and network orchestration, are

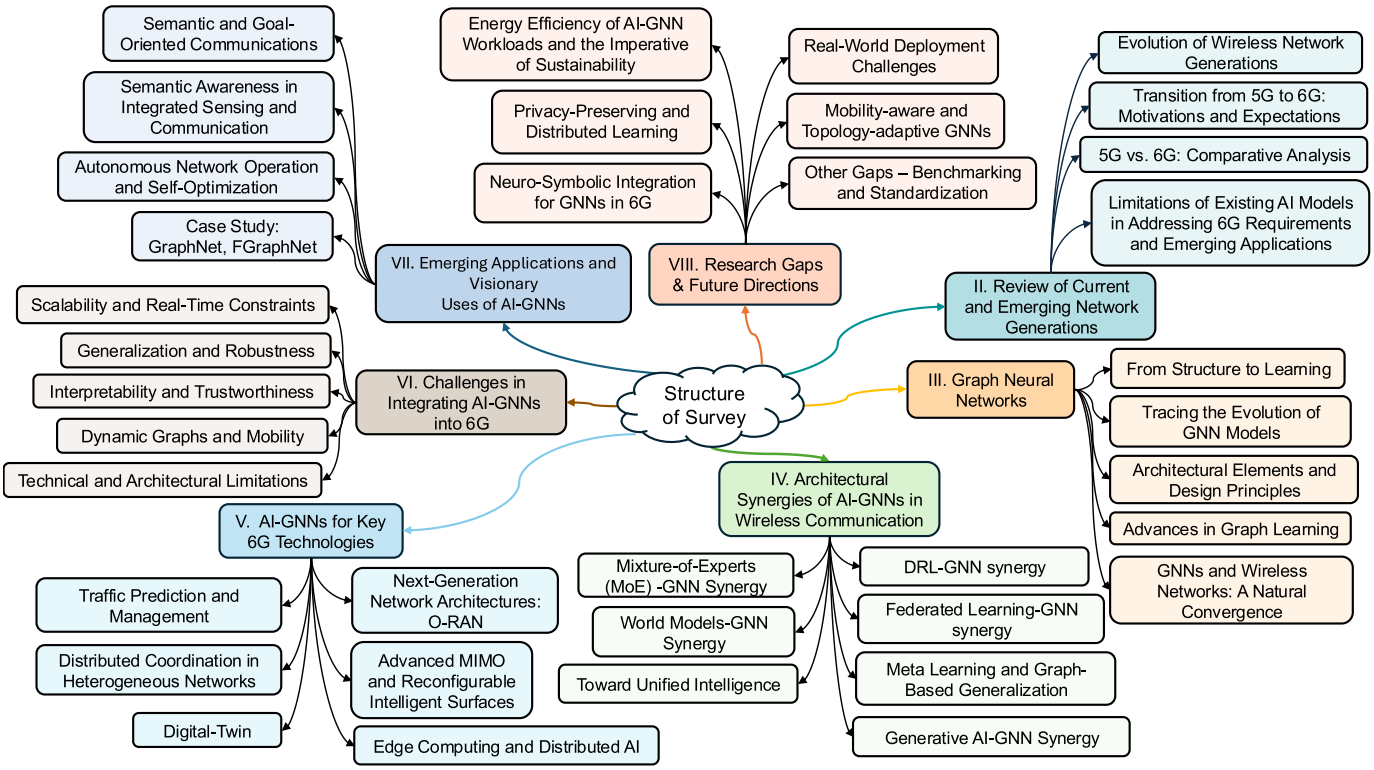


Fig. 1. Structure of this survey.

concisely captured in a comprehensive table at the end of this section. Section IV investigates the synergy between GNNs and other AI techniques, including RL, FL, meta-learning, generative AI, mixture-of-experts (MoE), and world models, and further highlights their collective potential toward unified, scalable, adaptive, and distributed intelligence in wireless systems. Section V discusses the role of AI–GNNs in enabling key 6G technologies, including Open Radio Access Network (O-RAN), advanced Multiple Input Multiple Output (MIMO) and Reconfigurable Intelligent Surface (RIS), edge computing and distributed AI, traffic prediction and management, distributed coordination in heterogeneous networks, and DTs. Section VI addresses the technical and architectural challenges of deploying AI–GNNs in real-world 6G scenarios, including scalability and real-time constraints, interpretability and trustworthiness, generalization and robustness, dynamic graphs and mobility, as well as broader technical and architectural limitations. Section VII explores emerging and visionary applications, including semantic and goal-oriented communications, semantic awareness in ISAC, autonomous network operation and self-optimization, and a case study on GraphNet and FGraphNet. Section VIII identifies critical research gaps and future directions, including the energy efficiency and sustainability of AI–GNN workloads, privacy-preserving and distributed learning, neuro-symbolic integration for GNNs in 6G, real-world deployment challenges, mobility-aware and topology-adaptive GNNs, and benchmarking and standardization. Finally, Section IX concludes the paper with a summary of insights and reflections on future research directions at the intersection of GNNs and intelligent wireless communication.

## II. Review of Current and Emerging Network Generations

### A. Evolution of Wireless Network Generations

Wireless communication systems have evolved through multiple generations over the past several decades, transitioning from simple voice-centric analog setups in the initial first-generation (1G) phase to advanced digital platforms handling rich data and media in fourth generation (4G). However, with growing demands for ultra-high-speed connectivity, low latency, massive device connectivity, and energy efficiency, the capabilities of previous generations have become insufficient, paving the way for 5G and eventually 6G networks [45], [46], [47].

### B. Transition From 5G to 6G: Motivations and Expectations

5G networks have introduced transformative technologies such as Millimeter Wave (mmWave), massive MIMO (mMIMO), network slicing, and Mobile Edge Computing (MEC). These advancements enable Enhanced Mobile Broadband (eMBB), Ultra Reliable Low Latency Communications (URLLCs), and Massive Machine Type Communications (mMTCs). Despite these improvements, 5G still faces critical challenges related to coverage, scalability, energy efficiency, and flexibility, which limit its ability to meet the long-term demands of future digital ecosystems [45], [46].

To address these limitations, the development of 6G networks has been initiated, driven by the vision of enabling ubiquitous global connectivity, terabit-per-second data rates, sub-millisecond latency, enhanced reliability, and autonomous

TABLE II
KEY FEATURES COMPARISON FOR 5G, 6G, AND AI/GNN CONTRIBUTIONS [17], [43], [45], [50]

| Key aspects | 5G | 6G | AI or GNN contributions |
|---|---|---|---|
| Applications | eMBB, URLLC, mMTC | Holographic-Type Communication, Tactile Internet, Intelligent Transport and Logistics, Intelligent and automated machines, virtual reality (VR), augmented reality (AR), XR | Modeling the complex, heterogeneous network interactions inherent in advanced applications. GNNs can enable proactive and context-aware orchestration, ensuring that diverse applications, from immersive XR to time-critical tactile internet services, receive the tailored resources and dynamic configuration needed for stringent performance and reliability requirements. |
| Peak data rate | ∼1–10Gbps | ∼1Tbps | Real-time AI-driven resource allocation and link adaptation to maximize throughput and approach theoretical peak rates under varying conditions. |
| End-to-End Latency | ∼1ms (in URLLC) | <0.5ms | Predictive scheduling and edge computing using AI for ultra-low latency such as caching and processing data closer to users, and GNN-based routing to find lowest-latency paths. |
| Reliability | $\sim 10^{-5}$ block error rate (high reliability) | Aiming for $\sim 10^{-9}$ (ultra-high reliability) | Proactive fault detection and failure mitigation. GNN models predict link/node failures and enable dynamic re-routing or resource reallocation. |
| Device Connectivity | Up to $10^6$ devices/km$^2$ | Expected to support up to $10^7$ devices/km$^2$ | Intelligent interference management and load balancing through AI to accommodate more devices. AI can learn traffic patterns and spatial interference maps, optimizing spectrum reuse and access scheduling in dense deployments. |
| Spectrum bands | Sub-6 GHz (low/mid bands) and mmWave up to ∼100 GHz (FR2) | Sub-THz/THz bands (100 GHz – 10 THz), and FR3 (7-24 GHz) | AI-based beamforming and adaptive modulation to mitigate propagation challenges at higher frequencies. ML can assist in real-time beam steering, blockage prediction, and frequency selection to mitigate the high path loss and directivity of THz links. |
| Coverage | Terrestrial with partial satellite integration | Seamless 3D coverage integrating terrestrial, aerial, and satellite links | AI-driven orchestration across heterogeneous networks. Handovers and multi-access connectivity between terrestrial and non-terrestrial platforms can be dynamically managed, ensuring users experience ubiquitous coverage and Quality of Service (QoS). |
| Energy efficiency | Incremental improvements over 4G | Targeting ultra-low energy per bit (on the order of ∼1 pJ/bit) for sustainable high-capacity networks | Energy-aware network protocols optimized via GNN insights. GNN-based optimization can identify energy-saving configurations in large-scale networks without compromising performance. |

network intelligence [45]. The International Telecommunication Union Radiocommunication Sector (ITU-R) has formally introduced the International Mobile Telecommunications (IMT)-2030 framework, outlining the vision, usage scenarios, and key capabilities for 6G mobile systems [48]. This framework highlights new service classes such as immersive communication (including XR and holographic media), AI-native networking, ISAC, and support for hyper-reliable, low-energy, and globally sustainable infrastructure [49]. A central theme of 6G, as emphasized by IMT-2030, is the seamless integration of AI and Machine Learning (ML) into the network fabric to enable real-time, context-aware, and self-optimizing operations [17].

These expectations impose stringent requirements on network design and signal processing capabilities, fostering a paradigm shift toward intelligent, adaptive, and collaborative network architectures that go beyond incremental improvements in 5G.

### C. 5G Vs. 6G: Comparative Analysis

A comparative analysis reveals that 6G networks are expected to substantially outperform 5G across multiple key performance indicators (KPIs), including data rate, latency, reliability, connectivity, spectral and energy efficiency, and the scope of intelligent functionalities. 6G seeks transformative improvements, leveraging advanced technologies such as sub-Terahertz (THz) and THz frequencies, extensive integration of aerial and space-based networks, and inherently integrates advanced AI and ML techniques. These include sophisticated GNNs, to enhance network performance, adaptability, and autonomy [17], [45]. Table II outlines a detailed comparison between 5G and 6G, along with specific contributions from GNN-Centric AI technologies.

The comparison in Table II highlights not only the quantitative enhancements of 6G over 5G but also demonstrates how GNN-Based AI technologies play a pivotal role in achieving the ambitious objectives of 6G. GNN-Centric AI Architectures are essential for managing complexity, enhancing efficiency, and enabling dynamic adaptability in future networks. Specifically, GNNs' capability to handle graph-structured data makes them uniquely suited for the challenges posed by next-generation wireless networks, offering significant benefits in resource management, predictive operations, and intelligent network optimization.

While Table II provides a comparative overview of the key differences between 5G and 6G, Figure 2 offers a conceptual visualization of how AI technologies, including ML, deep learning (DL), and GNNs, serve as the intelligence layer enabling the transition. The figure highlights how

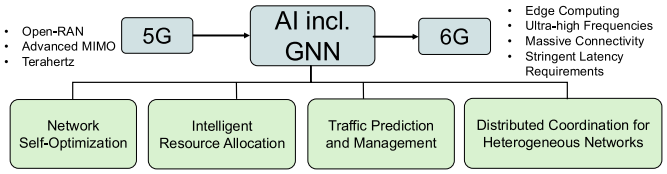


Fig. 2. AI as the intelligence layer bridging the gap between 5G and 6G wireless networks.

AI facilitates functions such as self-optimization, resource allocation, traffic management, and coordination in heterogeneous environments. The bullet points on the left side of the figure (O-RAN, advanced MIMO, THz) reflect technological advancements introduced in 5G, whereas the bullet points on the right (Edge Computing, Ultra-high Frequencies, Massive Connectivity, Stringent Latency Requirements) capture the emerging requirements and drivers that characterize 6G networks.

### *D. Limitations of Existing Ai Models in Addressing 6G Requirements and Emerging Applications*

6G and beyond systems must support massive connectivity, ultra-low latency, high reliability, native integration of communication–computation–sensing, and diverse service classes. While DL has been successfully applied to 5G design challenges, many existing models are built on assumptions misaligned with 6G realities.

First, most current AI models rely on fixed-size, Euclidean inputs and lack explicit representation of network topology. This limits their generalizability across varying network sizes and structures, often requiring retraining when users, base stations, or routing paths change. Such rigidity is ill-suited for dynamic 6G environments like ultra-dense, cell-free, or satellite–aerial–terrestrial networks, where graph structures evolve over time.

Second, these models typically optimize isolated functions (e.g., detection, adaptation, or scheduling) without capturing the cross-layer dependencies crucial for end-to-end performance. In tightly coupled 6G stacks, local decisions propagate globally through interference, congestion, or routing. Traditional architectures struggle to model such interactions, limiting their ability to support multi-service slicing and heterogeneous QoS requirements. Third, many AI models assume stationary environments, lacking mechanisms for adapting to dynamic graphs or non-stationary patterns. In 6G scenarios with rapid user mobility, handovers, and dynamic aerial or satellite node appearances, static models degrade significantly, jeopardizing ultra-reliable, low-latency communication.

Even existing GNN models, while inherently suited to relational data, face limitations when used in isolation. Many operate on single, static graph views and ignore temporal dynamics or domain constraints. Additionally, deeper GNNs struggle with scalability and training stability, particularly under edge resource constraints. These limitations highlight the need for hybrid GNN architectures that fuse graph-based inductive biases with complementary paradigms (e.g., attention, temporal modeling, optimization layers, and physics-aware modules) to deliver both expressivity and efficiency in 6G systems.

In summary, current AI models often assume static structure, stationarity, and modularity, making them ill-suited for the dynamic, multi-relational nature of 6G networks. This motivates a shift toward hybrid GNN-based solutions capable of (i) operating on evolving network graphs, (ii) integrating relational, temporal, and semantic views, and (iii) embedding domain knowledge and optimization principles. The remainder of this survey focuses on such architectures designed to meet the unique challenges of AI-native 6G systems.

## III. Graph Neural Networks

As discussed in earlier sections, AI lies at the core of the 6G vision for wireless connectivity, enabling networks that are not only fast and reliable but also adaptive, intelligent, and self-optimizing. Among emerging AI models, GNNs have gained particular prominence because of their ability to operate natively on graph-structured data, a form that fundamentally characterizes wireless network environments.

Building on the requirements and limitations identified in Section II, this section provides a comprehensive exploration of GNNs and their hybrid extensions as models inherently suited to network graphs, aligning with the structural and dynamic characteristics of 6G systems. We begin with the conceptual foundations and historical evolution of GNNs, then examine key architectures, learning paradigms, and recent innovations. In doing so, we highlight how GNNs differ from traditional neural networks, why they are uniquely positioned for 6G, and how they have evolved to handle increasingly complex, dynamic, and heterogeneous data environments.

### *A. From Structure to Learning*

GNNs are a class of neural architectures specifically designed to process *graph-structured data*, where entities are represented as nodes and relationships as edges. Such graph-based representations naturally map onto many real-world systems, including wireless communication networks, social networks, transportation grids, and biological structures [51]. Unlike conventional neural networks that operate on Euclidean domains (e.g., images or sequences), GNNs learn directly from *non-Euclidean* structures by exploiting relational and topological dependencies among entities [52].

Existing GNNs can be broadly divided into two categories: *spectral-based* models, which define graph convolutions via spectral filters of the graph Laplacian (capturing frequency-domain characteristics of signals on a graph), and *spatial-based* models, which operate in the vertex domain and propagate information through localized message passing among neighboring nodes.

Among spatial approaches, the message passing neural network (MPNN) framework [53] has emerged as a unifying abstraction. In an MPNN, each node iteratively updates its hidden state by exchanging information (messages) with its neighbors through three steps: message computation,

aggregation, and update. The general form of this process at the $k$-th layer can be expressed as:

$$\begin{aligned}\mathbf{m}_i^{(k)} &= \mathcal{A}^{(k)}\Big(\{\mathcal{M}^{(k)}(\mathbf{h}_i^{(k-1)}, \mathbf{h}_j^{(k-1)}, \mathbf{e}_{ij}) \mid j \in \mathcal{N}(i)\}\Big),\\ \mathbf{h}_i^{(k)} &= \mathcal{U}^{(k)}\big(\mathbf{h}_i^{(k-1)}, \mathbf{m}_i^{(k)}\big),\\ \mathcal{M}^{(k)} &: \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d_e} \to \mathbb{R}^{d_m}.\end{aligned} \tag{1}$$

Here, $\mathcal{N}(i)$ is the set of neighbors of the node $v_i$, $\mathbf{m}_i^{(k)} \in \mathbb{R}^{d_m}$ and $\mathbf{h}_i^{(k)} \in \mathbb{R}^d$ denote the aggregated message and the node embedding of $v_i$ at layer $k$, respectively. The edge attributes between nodes $v_i$ and $v_j$ are represented by $\mathbf{e}_{ij} \in \mathbb{R}^{d_e}$, which denotes the edge feature vector associated with edge $(v_i, v_j)$. The message function $\mathcal{M}^{(k)}$ computes pairwise messages based on the representations of adjacent nodes and their connecting edge features, where $d$, $d_e$, and $d_m$ denote the dimensions of node embeddings, edge features, and message vectors, respectively. The aggregation function $\mathcal{A}^{(k)}$ is a permutation-invariant operator (e.g., summation, mean pooling, or attention) that combines incoming messages from the neighborhood $\mathcal{N}(i)$. Finally, the update function $\mathcal{U}^{(k)}$ integrates the aggregated message to produce the updated node embedding. While some models employ learnable edge embedding functions to capture relation-specific information, others directly utilize raw edge attributes when available.

A prominent instantiation of this general framework is the Graph Convolutional Network (GCN) proposed by Kipf and Welling [54], which combines message aggregation and node-wise update into a single normalized propagation and transformation operation:

$$H^{(k+1)} = \sigma\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}H^{(k)}W^{(k)}\right), \tag{2}$$

where $H^{(k)} \in \mathbb{R}^{n\times d}$ is the matrix of node embeddings at layer $k$, $W^{(k)}$ is a learnable weight matrix, $\tilde{A} = A + I_n$ denotes the adjacency matrix augmented with self-loops, $\tilde{D}$ is the corresponding degree matrix, and $\sigma(\cdot)$ is a nonlinear activation function (e.g., ReLU). This formulation couples a linear transformation with a normalized neighborhood aggregation operation, and can be interpreted as a first-order approximation of spectral graph convolution. In wireless communication graphs, nodes may represent BSs or UEs, while edges encode interference or coupling relationships. The normalized propagation helps stabilize training under heterogeneous node degrees and network densities.

By explicitly incorporating graph topology into the learning process, GNNs offer significant advantages for modeling the complex interactions inherent in wireless networks. In particular, their permutation invariance to node ordering and their ability to encode relational and topological priors make them well suited for learning scalable and adaptive communication, scheduling, and resource management policies in large-scale, dynamic, and heterogeneous systems.

### B. Tracing the Evolution of GNN Models

The conceptual foundations of GNNs can be traced back to message-passing algorithms in probabilistic graphical models (such as belief propagation on factor graphs) that emerged in the 1990s [55], [56]. These earlier techniques introduced the idea of iteratively exchanging information between connected nodes, a principle that later inspired the design of trainable graph neural architectures [57]. Formal definitions and prototypes of graph neural networks began taking shape in the mid-2000s, with foundational works by Gori et al. [58], Scarselli et al. [59], and others [60] who proposed recursive neighborhood-based learning algorithms on graphs. These early GNNs focused on learning node representations through iterative aggregation of neighbor information, laying the groundwork for modern graph-based learning systems.

Subsequent years witnessed rapid innovation in graph neural network research. The development of the Graph Attention Network (GAT)s that leverage attention mechanisms for neighborhood weighting [61], marked a new era of scalable deep learning on graphs. Around the same time, techniques for model simplification and efficiency began to address the scalability of GNNs in large graphs [62]. New specialized architectures also emerged; for example, Eigen-GNN [63] improved the adaptability of graph models by combining spectral graph properties with learnable parameters, demonstrating enhanced performance across diverse tasks.

As the field matured, GNNs were extended into privacy-preserving and secure learning architectures for distributed data [64], [65]. The introduction of models like the Relational Graph Convolutional Network (R-GCN) further expanded GNN capabilities by incorporating relation-specific message functions to handle multi-relational knowledge graphs [66].

In summary, while the earliest GNN models were largely static and defined by recursive graph traversals, contemporary GNNs encompass a variety of spectral and spatial designs that employ differentiable, end-to-end trainable message-passing mechanisms influenced by modern deep learning practices [67].

### C. Architectural Elements and Design Principles

At the heart of every GNN are three fundamental components: *node embeddings*, which represent each node's features or latent state; *message functions*, which define how nodes communicate and exchange information with their neighbors; and *aggregation functions*, which combine the messages received from neighbors before updating the node's own state.

In more advanced GNN variants, one can also define embedding functions for edges to incorporate relationship-specific information between nodes (as is in models like R-GCN). If rich feature descriptors for nodes and edges are directly available from data, explicit learnable embedding layers can sometimes be omitted.

Modern GNN architectures often include additional enhancements such as multi-head attention mechanisms, residual or skip connections, normalization layers, and adaptive activation functions. These design improvements stabilize training and enable scaling to deep or large graphs, which is crucial for noisy, large-scale, and heterogeneous networks.

### D. Advances in Graph Learning

While traditional GNNs have demonstrated strong performance on static, well-defined graph structures, real-world

systems, especially wireless networks, are inherently dynamic, heterogeneous, and often partially observable. As a result, a major frontier in GNN research focuses on *dynamic graph learning*, where the structure, features, and connectivity of the graph evolve over time. This shift is particularly crucial for 6G applications, where network topologies change due to user mobility, spectrum reallocation, link failures, and dynamic service demands.

To address these challenges, several *temporal and structure-aware GNN frameworks* have been proposed. One such model is *GraphSense* [68], which introduces a self-aware dynamic structure learning paradigm that integrates Recurrent Neural Network (RNN) backbone with a self-sensing neighborhood aggregation module. This dual-module architecture enables GNNs to track temporal variations and autonomously adapt to evolving graph topologies without relying on full graph visibility, a critical requirement for scalable and distributed wireless intelligence. Similarly, the *EvolveGCN* framework [69] offers a solution for temporal graphs by dynamically updating the GNN weights through recurrent models (e.g., Gated Recurrent Unit (GRU)/Long Short Term Memory networks (LSTM)) as new graph snapshots are introduced. EvolveGCN enables temporal modeling at both the node and edge levels, making it suitable for time-series forecasting in mobile communication environments.

Beyond temporal modeling, another important direction is the rise of *self-supervised graph learning*. Given the scarcity of labeled data, self-supervised approaches such as *Graph Contrastive Learning (GCL)* have become increasingly valuable. These methods learn meaningful graph embeddings by maximizing the agreement between augmented views of the same node or subgraph. However, traditional contrastive learning methods often capture spurious correlations that hinder generalization.

To address this, recent work by Mo et al. [70] introduces a *causality-aware graph contrastive learning* framework that disentangles causal from non-causal factors using spectral graph augmentation and customized invariance-independence training objectives. This model improves robustness and interpretability, which is essential for critical applications such as anomaly detection and intelligent control in wireless networks. Similarly, the authors in [71] propose a novel Counterfactual Contrastive Learning framework (CCL-Gn) to mitigate Endogenous and Task-oriented Spurious Correlations (ETSC) in GNNs, particularly in node-level tasks. Their method automatically separates spurious subgraphs from causally correlated subgraphs, improving GNN robustness in both in-distribution and out-of-distribution scenarios. This framework can be particularly relevant in real-world wireless communication applications, where data inconsistencies and unreliable patterns can significantly impact performance.

Additional efforts in dynamic graph representation learning include Temporal Graph Attention Networks (TGAT) [72], which combines temporal encoding with graph attention to model asynchronous events over time; Dynamic Self-Attention Network (DySAT) [73], which learns both structural and temporal self-attention to capture time-varying graph dependencies; and Temporal GCN (T-GCN) [74], which integrates GCNs with GRUs to model spatial–temporal dependencies, originally developed for traffic forecasting but adaptable to latency prediction and traffic shaping in wireless systems. Another promising direction is the application of meta-learning and continual learning in GNNs for dynamic environments. For example, Meta-GNNs [75] are designed to rapidly adapt to new topologies through a meta-optimization objective, making them particularly suitable for non-stationary, highly mobile wireless scenarios.

In parallel, the growing emphasis on explainability and trustworthiness has driven research into interpretable and causal graph learning. Approaches such as GNNExplainer [76] and PGM-Explainer [77] offer post hoc interpretability by revealing important substructures or probabilistic dependencies that influence model predictions. Within 6G contexts, these methods can enhance transparency by elucidating decision processes in autonomous resource allocation, explaining anomalies in distributed systems, and justifying spectrum-sharing strategies in AI-managed networks.

These advancements demonstrate that GNNs are no longer limited to learning from fixed graph snapshots. Instead, they are evolving into dynamic, self-supervised, and causally informed learning agents, capable of reasoning in complex, high-variance environments, precisely the kind of intelligence required for next-generation wireless systems. As the field progresses, integrating these graph learning advances into communication network design will be instrumental in fulfilling the vision of 6G: *adaptive, autonomous, and AI-native connectivity*.

### *E. GNNs and Wireless Networks: A Natural Convergence*

Wireless networks are inherently dynamic, heterogeneous, and graph-structured systems, where entities such as base stations, users, or IoT devices can be naturally modeled as nodes, and their interactions, ranging from signal strength to latency or interference, are represented as edges. This natural graph abstraction stands in contrast to the Euclidean assumptions of traditional AI models like convolutional or recurrent neural networks, which are ill-suited for capturing the irregular and evolving dependencies found in communication systems. In contrast, GNNs are explicitly designed for learning on graph-structured data, enabling them to model the relational, topological, and context-aware nature of modern and future wireless environments [39].

Core to their utility in communication systems is the *message-passing paradigm*, which allows each node to iteratively aggregate information from its neighbors, making GNNs inherently scalable and well-suited for decentralized inference in distributed edge deployments [103]. This mechanism also grants GNNs *topology-awareness* and the ability to generalize across different network sizes and structures, making them robust under mobility, varying density, and topological shifts. Furthermore, GNNs can seamlessly integrate with RL and FL, enabling intelligent, privacy-preserving, and adaptive control in highly dynamic environments.

Several specialized GNN architectures have emerged to support diverse communication tasks. *GCNs* use spectral or spatial convolutions for topology-aware feature propagation,

while *GATs* assign adaptive weights to neighbors during aggregation based on learned attention coefficients. *MPNNs* generalize this paradigm and are widely applied in tasks such as scheduling, routing, and power control. More recent models, like *SAGNN* [134], leverage subgraph encoding for structural sensitivity, while the *SACGNN* framework captures dual communication–sensing interactions for edge intelligence.

Beyond conventional GNNs, new graph-based modeling extensions are gaining traction. The *multi-functional time-expanded graph (MF-TEG)* [135] framework enables optimization across joint user association and service routing in space–air–ground networks. Meanwhile, *Hypergraph Neural Networks (HGNNs)* generalize standard pairwise GNNs by introducing hyperedges that connect multiple nodes simultaneously. This allows richer modeling of group-based interactions such as coordinated multi-point transmission (CoMP), UE clustering, and joint beamforming [136], thereby capturing higher-order dependencies essential for distributed coordination in cell-free 6G networks.

Taken together, these developments illustrate why GNNs have become an increasingly important learning tool for wireless communications. To provide intuition on this evolution, Table III presents a *representative pre-synergy overview* of GNN-based applications across wireless communication tasks. Rather than aiming for exhaustive coverage, the table illustrates the breadth of existing approaches and highlights key technical directions in which GNNs have shown clear promise (including resource management, physical-layer processing, network intelligence, and emerging 6G enablers) while emphasizing representative technical distinctions within each sub-application. A more systematic and comprehensive taxonomy, centered on the broader AI–GNN synergy perspective of this survey, is developed in the subsequent sections.

## IV. Architectural Synergies of AI–GNNs in Wireless Communication

Recent advances in AI and graph learning have revealed promising synergies between GNNs and advanced AI paradigms such as DRL, FL, meta-learning, and XAI. Rather than serving as a universal solution, GNNs provide a structured modeling framework that is particularly effective when relational dependencies, topology, and distributed interactions play a central role.

In this context, hybrid AI–GNN models offer a flexible approach to combining graph-based representation learning with complementary AI techniques, enabling more adaptive, scalable, and context-aware decision-making in wireless systems. This section presents a structured analysis of these synergies, highlighting how such integrations can enhance specific classes of wireless applications, particularly those involving distributed control, partial observability, and strong inter-node coupling.

### A. DRL-GNN Synergy

The integration of GNNs with DRL enables topology-aware sequential decision-making in wireless networks, where GNNs capture relational structure and DRL handles uncertainty and dynamics. This combination is particularly effective in scenarios with strong interdependencies, but its advantages may diminish when such structure is weak or can be approximated by simpler models.

Figure 3 presents a high-level architecture of this synergy comprising two complementary modeling perspectives. From a first perspective, each agent is represented as a node in a graph, and edges denote communication or coordination links between agents. This interaction graph enables agents to exchange information through GNN-based message passing, facilitating cooperative policy learning under partial observability.

From a second perspective, the wireless environment itself can be represented as a graph, where nodes correspond to network components such as base stations, UAVs, or users, and edges capture physical or logical dependencies, including interference, connectivity, or task coupling. This representation allows GNNs to extract relational features that can enrich the DRL state representation. The DRL agent then operates on this structured representation to evaluate actions and their system-level impact.

This hybrid formulation is particularly relevant for multi-objective and distributed optimization problems, where local decisions (e.g., power control or routing) influence globally coupled performance metrics such as throughput, latency, and energy efficiency.

*1) Multi-Agent Coordination and Graph-Based DRL:* MARL has emerged as a key enabler for distributed decision-making in wireless systems. However, conventional MARL methods often face challenges related to scalability, non-stationarity, and communication overhead. To address these issues, recent research has explored the integration of GNNs into MARL frameworks, leveraging their ability to model interaction topologies among agents and support efficient information exchange [137], [138], [139], [140].

A prominent line of work incorporates GNN-based communication mechanisms within MARL to better capture dynamic agent relationships. For instance, GNNComm-MARL integrates GNNs for scalable communication in partially observable environments, while dynamic graph construction enhances generalization across varying topologies. Permutation-invariant message-passing architectures further enable efficient aggregation of neighborhood information, improving coordination in scenarios such as UAV swarms and cell-free MIMO [141], [142]. These studies suggest that graph structures can provide useful inductive biases for decentralized coordination, particularly in large-scale and dynamic networks.

Building on this idea, several works apply GNN-enhanced MARL to specific wireless optimization tasks. In [143], GNN-assisted MARL improves spectrum allocation by explicitly modeling spatial dependencies among agents. Similarly, [144] proposes a heterogeneous attention graph-based multi-agent DRL algorithm that jointly optimizes access and resource allocation, improving coordination among heterogeneous agents and enhancing overall system performance.

Further extending these approaches, [145] introduces graph reinforcement learning (GRL) for joint routing and resource allocation in wireless ad hoc networks. Their decentralized,

TABLE III
REPRESENTATIVE PRE-SYNERGY OVERVIEW OF GNN APPLICATIONS IN WIRELESS COMMUNICATION TASKS

| Task Group | Sub-Application | Representative Distinction | Representative Works |
|---|---|---|---|
| **Resource Management** | Power Control & Resource Allocation | These works focus on topology-aware power control and generic radio resource allocation over interference graphs. Their distinction lies in learning scalable allocation policies that generalize across network sizes and topologies, including domain-informed designs such as unrolled optimization. | [78]–[82] |
| | Fairness/QoS Allocation | Unlike pure throughput maximization, these works explicitly model user fairness, QoS, or utility trade-offs. GNN-based policies exploit relational user–BS structure to balance performance across heterogeneous users under practical constraints. | [83] |
| | Link Scheduling | These works formulate interference-aware scheduling as a graph problem, often related to MWIS-style selection. Their main distinction is the use of message passing to learn contention-aware activation policies beyond greedy heuristics. | [84], [85] |
| | Task Scheduling & Offloading | In contrast to radio-only optimization, these works learn graph-aware embeddings of tasks, servers, or network states for latency-sensitive scheduling and offloading in heterogeneous IoT/edge systems. | [86]–[89] |
| **Physical Layer Processing** | Beamforming | These works include fully digital beamforming designs for MU-MISO, MISO interference channels, and high-mobility scenarios such as high-speed railway systems. Their main distinction is the use of message passing to capture inter-user/inter-link coupling while enabling low-complexity online inference. | [90]–[93] |
| | NOMA Power Allocation | These works specialize GNNs for non-orthogonal multiple access, learning graph-aware power allocation policies that remain scalable and efficient under strong interference coupling and user superposition constraints. | [94] |
| | Precoding & Joint Scheduling/Precoding | These works target spatial signal design, often jointly with user scheduling, and differ from generic resource allocation by directly optimizing transmit-side PHY variables under channel-dependent coupling. | [87], [95]–[99] |
| | Hybrid Beamforming | These works explicitly model analog/digital beamforming structures, including homogeneous and heterogeneous graph constructions for mmWave systems, thereby improving scalability and adaptability in large antenna arrays. | [100], [101] |
| | MIMO Detection | GNN-based detectors learn from interference or factor-graph structures and differ from classical model-based detectors by generalizing across antenna/user configurations with lower mismatch sensitivity. | [102]–[104] |
| | Channel Decoding | These works apply learned message passing over code/factor graphs and differ in the targeted code families and decoding strategies, often improving convergence or adaptability over classical BP-based decoders. | [105]–[109] |
| **Security & Reliability** | Intrusion & Cybersecurity | These works model structural dependencies in network, vehicular, or protocol traffic graphs. Their distinction lies in combining topology-aware anomaly signatures with limited-label learning for robust intrusion and attack detection. | [110]–[114] |
| | Encrypted Traffic Classification | Compared with flow-statistics-based methods, these works construct packet/flow interaction graphs to capture multi-scale communication patterns, improving classification accuracy under encryption and traffic obfuscation. | [112], [113] |
| | Anomaly/Fault Detection | These studies differ in whether they emphasize spatio-temporal KPI correlation, topology-aware failure prediction, or digital-twin-assisted diagnosis, but all use GNNs to exploit relational dependencies for early fault identification. | [115], [116] |
| **Network Management** | Routing, Scheduling & Control | These works learn scalable local or distributed control policies over dynamic topologies. Their distinction lies in targeting routing, network control, and topology-aware management decisions under varying traffic and resource conditions. | [86], [117]–[119] |
| | Graph Representation & Knowledge Modeling | Rather than directly optimizing radio variables, these works construct compact graph embeddings of network state, context, or semantic knowledge, enabling efficient information exchange and semantic-aware communications. | [120], [121] |

TABLE III
*(Continued.)* REPRESENTATIVE PRE-SYNERGY OVERVIEW OF GNN APPLICATIONS IN WIRELESS COMMUNICATION TASKS

| | | | |
|---|---|---|---|
| **Emerging 6G Enablers** | RIS / Flexible & Pinching Antennas | These works target reconfigurable propagation and emerging antenna paradigms (e.g., RIS and pinching antennas). Their distinction lies in modeling interaction structures such as element–user coupling, joint transceiver–RIS–relay optimization, and antenna–channel coupling, enabling scalable learning of phase shifts and beamforming. | [122]–[124] |
| | Cell-Free / Distributed MIMO | These works model distributed AP–UE interactions as graphs and distinguish themselves by enabling scalable coordination, power control, and inference across dense distributed MIMO architectures. | [93], [101] |
| | Integrated Sensing & Communication | These works differ from communication-only optimization by jointly balancing sensing and communication objectives, often using structured graph modeling or graph pruning to reduce the complexity of ISAC design. | [125], [126] |
| | Localization & Radio Map Reconstruction | These works exploit spatial signal graphs for localization and radio map reconstruction. Their distinction lies in learning geometry-aware signal relationships from sparse CSI or spectrum observations to improve positioning accuracy and propagation map completion. | [127]–[131] |
| | Satellite / SAGIN Networks | These works use distributed GNN policies for non-terrestrial and space–air–ground integrated networks, enabling adaptive routing, handover, and load balancing through locality-aware decision making. | [132], [133] |

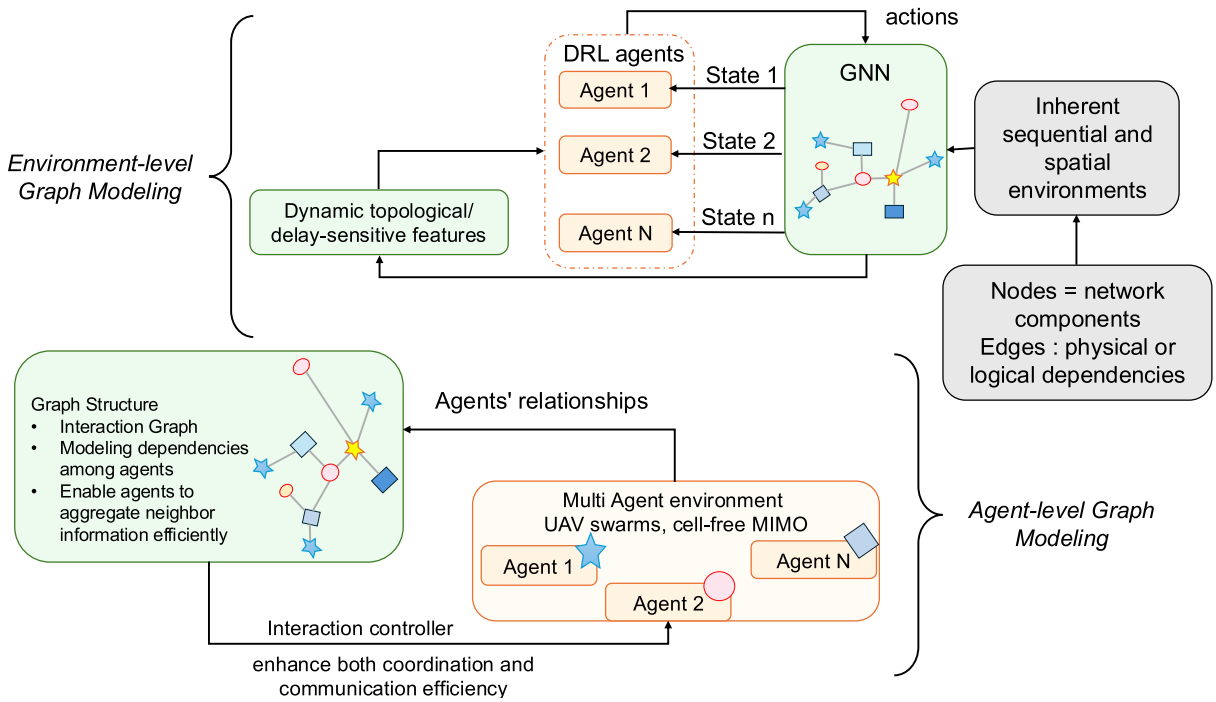


Fig. 3. DRL-GNN synergy: agent-level and environment-level graph modelling perspectives.

hop-by-hop framework leverages graph representations to capture irregular topologies and interference patterns, using a GNN policy network to process variable-sized inputs efficiently. The method demonstrates improved performance compared to heuristic and conventional DRL baselines in terms of throughput and connectivity.

Finally, [146] reduces the computational complexity of DRL-based resource allocation by incorporating relational priors such as topology awareness and permutation invariance. By representing the system state as a graph and adopting a graph-based Deep Deterministic Policy Gradient (DDPG), the approach achieves comparable performance with faster convergence and reduced overhead relative to standard DDPG methods.

*2) GNNs for Structured Decision-Making in Routing and Trajectory Planning:* Applications such as IoT routing, UAV trajectory planning, and network slicing are inherently sequential and spatial, making them ideal for DRL frameworks enhanced by GNNs. In consumer-centric IoT, models like SAG use GNNs to extract dynamic topological features and integrate attention mechanisms with Soft Actor-Critic (SAC) for intelligent routing decisions [147]. Hybrid models combining GCNs with DRL significantly improve routing convergence and resource efficiency [148]. In UAV systems, GNNs enable energy-efficient and robust 3D localization and multi-agent trajectory design, handling GNSS-denied and disaster scenarios with improved coordination and system utility [149], [150].

*3) GNNs in Satellite and Network Slicing Optimization:* In satellite networks, particularly LEO mega-constellations, GNN-based DRL algorithms such as GRL-RR and GRLR model spatially distributed nodes and inter-satellite links for resilient and adaptive routing [151], [152]. GraphSAGE and other GNN variants are used to capture topological dynamics and feed delay-sensitive features into DRL agents [153]. In network slicing, FGRCNN and GAT-based frameworks dynamically orchestrate resources across time-varying graph structures, improving slice embedding accuracy, service quality, and embedding efficiency [154], [155].

*4) Caching, Scheduling, and Task Offloading:* GNN-based DRL frameworks are also applied in cooperative caching and content distribution, with GATs and dueling DQNs improving video selection and edge caching latency [156], [157]. In task offloading, works like EMDTORA and GCP-DDPG utilize GNNs to model task dependency graphs and inform DRL decisions under dynamic network loads [158], [159]. Other efforts use GNNs to learn topological embeddings for multi-agent optimization of task offloading and service utility [160], [161].

*5) Spectrum Access, Scheduling, and Network Orchestration:* For spectrum allocation, interference-aware GRL frameworks model device interactions as hypergraphs or dynamic topologies [162], [163]. In TSN and UAV scheduling, GATs and GCNs capture spatial-temporal relationships for delay-aware resource management [164], [165]. Models like GRLinQ and GROM apply GRL to optimize link scheduling, routing, and virtual network function (VNF) placement in dynamic topologies [166], [167]. In this direction, [145], [146] further enhance dynamic spectrum access and reduce DRL complexity using topology-aware policy design, demonstrating strong throughput and latency performance in decentralized and multi-hop environments.

### B. Federated Learning-GNN Synergy

As wireless networks evolve toward decentralized and privacy-aware architectures, FL has emerged as an effective framework for distributed learning. FL enables edge devices to collaboratively train global models without exchanging raw data, thereby preserving user privacy and reducing reliance on centralized data collection. However, conventional FL faces several practical challenges, including communication overhead, device heterogeneity, and scalability limitations, which become more pronounced in dynamic and resource-constrained wireless environments.

Recent studies have explored the integration of GNNs into FL frameworks, leading to what is commonly referred to as graph-based federated learning (FGL). In this context, GNNs provide a structured way to model relationships among devices, data distributions, or network topology, enabling more informed aggregation and collaboration strategies. In particular, graph representations can help capture non-IID data correlations, spatial dependencies, and connectivity patterns, which are not explicitly considered in conventional FL. As a result, FGL can improve coordination efficiency and model robustness in scenarios where relational structure plays a significant role. However, these benefits are application-dependent and may introduce additional complexity in graph construction and maintenance.

Figure 4 illustrates a representative architecture of an FGL system. Each device trains a local GNN on its private data, where nodes and edges encode device-specific features and relational dependencies. Local model updates are then transmitted through a secure sharing layer, ensuring privacy-preserving communication. A central server aggregates these updates to construct a global model and coordinate the learning process. From a system perspective, FGL can also support adaptive collaboration strategies, synchronization policies, and resource-aware learning, for example by adjusting aggregation frequency to meet latency constraints or by balancing communication and computation in resource-limited environments.

*1) Privacy, Security, and Efficiency:* Several recent studies focus on enhancing privacy and efficiency in GNN-based FL frameworks. Techniques such as FedGAT and differentially private GNNs are designed to mitigate inference leakage risks in wireless networks by embedding privacy guarantees during model updates [168], [169]. Lightweight frameworks, including LW-FGL and JC3ES, introduce cost-aware communication bottlenecks and dynamic collaboration strategies tailored for time-sensitive applications like UAV networks and semi-supervised edge learning [170], [171]. Additionally, energy-efficient FGL methods exploit graph-driven feature and neighbor selection to minimize power consumption and reduce computational overhead during training and transmission [172], [173].

In FGL, data security and user privacy are strengthened by graph-aware defenses layered on standard FL protections: clients keep raw nodes/edges local while sharing only masked statistics or model deltas; cryptographic secure aggregation makes the server learn only sums, thwarting single-client gradient/embedding inspection and improving robustness when combined with backdoor-detectability checks [174], [175]; differential privacy (central or local) perturbs structure-revealing signals (e.g., neighbor embeddings, message-passing messages) with calibrated noise to bound graph/attribute leakage [176]; and graph-specific designs restrict what structural evidence is exposed, e.g., training on privacy-preserving ego-graphs or subgraph views and exchanging obfuscated summaries rather than adjacencies, which mitigates topology disclosure and supports heterogeneous clients [177], [178]. Such mechanisms end-to-end for graphs, covering attack surfaces (membership/link/attribute inference) and reporting that, with careful noise/aggregation choices, FGL can retain strong utility on tasks like recommendation and traffic prediction while materially reducing leakage in decentralized settings [179].

*2) Heterogeneity, Scalability, and Orchestration:* To handle data and resource heterogeneity, hierarchical FGL systems incorporate graph-based orchestration techniques that jointly optimize client selection, resource allocation, and scheduling [123], [180]. In large-scale deployments, techniques such as GNN-enhanced over-the-air computation (AirComp) improve scalability and robustness by enabling direct model aggregation in the wireless medium, reducing latency and communication costs [181]. These advancements make FGL a

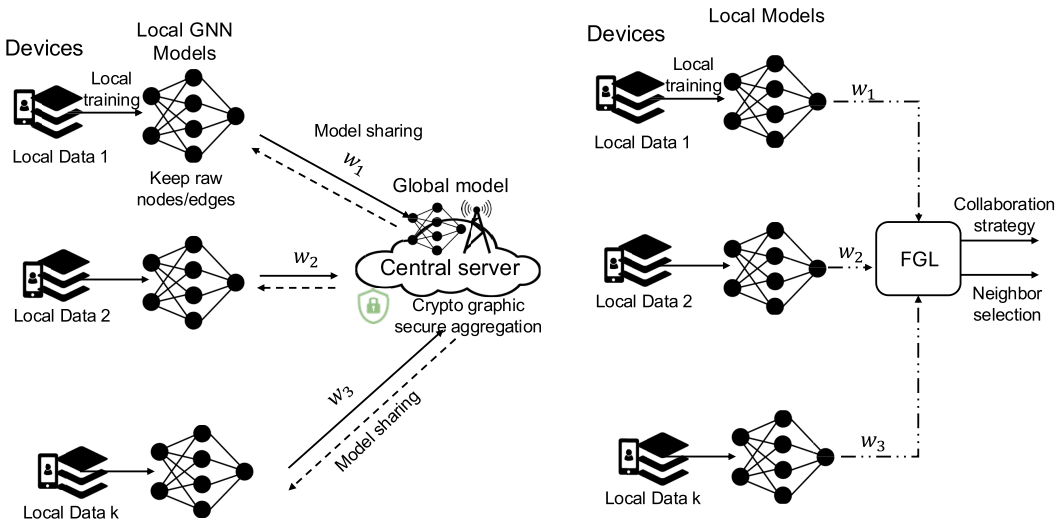


Fig. 4. Synergy between FL and GNNs. GNNs operate either as local, structure-aware models within federated learning or as coordination layers that adapt communication topology and collaboration strategies to optimize latency, energy efficiency, and task performance.

viable solution for next-generation networks requiring secure, efficient, and scalable learning across distributed and dynamic environments. FL enhances privacy and decentralization in AI models but faces challenges in communication overhead, heterogeneity, and limited scalability. The GNN-FL synergy enables structure-aware and communication-efficient collaboration among edge devices [181], [182], [183].

### C. Meta Learning and Graph-Based Generalization

Graph meta-learning (GML) provides a flexible framework for adapting models under non-stationary conditions and limited data availability, which are common in wireless environments. By combining GNN-based relational representations with meta-learning strategies, these approaches aim to improve generalization across varying network topologies and task distributions.

For instance, works such as MetaGraphLoc and hybrid-task meta-learning demonstrate improved adaptability for localization and bandwidth allocation across heterogeneous environments [75], [184]. Similarly, task offloading in SAGIN and dynamic MEC scenarios has been explored using GNN-based meta-models to enable rapid policy adaptation under changing system conditions [185]. In [186], a meta-learned GNN for power control employs a Random Edge GNN (REGNN) combined with modular meta-learning to facilitate efficient fine-tuning across different network topologies with limited training data.

Overall, these studies suggest that integrating GNNs with meta-learning can improve adaptability in scenarios characterized by structural variability and data scarcity. However, the effectiveness of such approaches depends on the availability of representative meta-training tasks and may involve additional training complexity compared to conventional learning frameworks.

### D. Generative Ai–GNN Synergy

Recent studies have explored the integration of GNNs with generative AI techniques, including VAEs, GANs, and denoising diffusion models, to address wireless networking problems in a data-driven manner [187]. In this context, GNNs are used to encode relational structure in wireless data, enabling generative models to produce synthetic network states or candidate solutions that reflect underlying dependencies and uncertainties. Such approaches are particularly relevant in scenarios where data is scarce, expensive to obtain, or highly variable.

Generative diffusion models (GDMs), in particular, have attracted attention due to their robustness and stable training behavior [188]. These models have been applied across multiple layers of wireless systems, from physical-layer signal modeling to higher-layer network optimization, to generate representative data samples for system design and evaluation [189]. When combined with GNNs, diffusion models can incorporate topology-aware information, enabling the generation of structured outputs such as resource allocation patterns or network configurations. However, these benefits come at the cost of increased computational complexity and may require careful design to ensure scalability in large networks.

One example is the use of GNN-parameterized diffusion models for wireless resource optimization [190]. In this approach, a diffusion model is trained to approximate an expert policy by generating feasible resource allocation patterns from noise, while the GNN backbone encodes graph-structured relationships such as interference or user association [191]. This combination allows the model to produce candidate solutions that are consistent with network structure and can generalize across different topologies. Nevertheless, performance depends on the quality of the training data and the underlying expert policy.

More broadly, graph-based generative models have been proposed to produce graph-structured data representations for wireless systems [192]. These approaches adapt generative processes to structured domains by encoding network states as graph signals and sampling new configurations using GNNs or graph transformers. While such methods offer a flexible framework for simulation and scenario generation, their practical deployment in real-time wireless systems remains

an open challenge due to latency, scalability, and reliability considerations.

### E. Mixture-of-Experts (MoE)-GNN Synergy

The integration of MoE with GNNs offers a principled approach to scaling graph-based learning under the heterogeneity and dynamicity of wireless systems. Unlike monolithic GNNs, MoE architectures introduce multiple expert subnetworks together with a routing mechanism that selectively activates only a subset of experts per input, enabling conditional computation and improved model capacity [193], [194].

A key challenge in MoE–GNN design is determining *the granularity of expert routing*. Recent work on Graph Mixture-of-Experts (GMoE) demonstrates that routing can be performed at the node level, allowing each node to dynamically select aggregation experts based on its local structural context [195]. This is particularly relevant for wireless networks, where heterogeneity is inherently localized. For example, nodes corresponding to users at the cell edge, dense interference regions, or highly mobile clusters may require fundamentally different aggregation behaviors compared to more stable or sparsely connected regions. By enabling node-wise expert selection, MoE–GNN architectures can adapt to such spatial and temporal variations more effectively than globally shared models.

From a modeling perspective, MoE–GNN enables *expert specialization over graph structural diversity*. In GMoE, different experts are designed with distinct inductive biases, such as varying aggregation ranges (e.g., hop-1 versus hop-2 message passing), allowing the model to capture both local and long-range dependencies within a unified architecture [196]. This is particularly important in wireless systems, where some tasks (e.g., local interference management) require short-range interactions, while others (e.g., coordinated beamforming or cell-free processing) depend on long-range coupling. Such diversity-aware modeling aligns naturally with the multi-scale and heterogeneous nature of wireless graphs.

However, the effectiveness of MoE–GNN critically depends on the design of the *routing mechanism*. In practice, routing is typically implemented via a gating network that maps node embeddings to expert selection scores, often using top-$k$ sparse activation strategies [194]. In dynamic wireless environments, this gating function must adapt not only to static topology features but also to time-varying conditions such as traffic demand, mobility patterns, and interference fluctuations. This introduces additional challenges in maintaining stable and consistent expert assignments over time.

Beyond routing, *expert load balancing* emerges as a fundamental challenge. As shown in recent MoE and GMoE studies, naive training may lead to expert collapse, where only a small subset of experts is consistently selected due to self-reinforcing feedback in the routing mechanism [197]. To address this, explicit balancing objectives are introduced, including importance-based regularization to equalize expert contribution and load-based constraints to ensure uniform selection probability across experts. These mechanisms are essential to prevent over-utilization of certain experts and under-training of others.

In wireless networks, load balancing is further complicated by non-uniform and time-varying graph distributions. For instance, hotspot formation, bursty traffic, or mobility-induced topology changes may cause large portions of the graph to be routed to the same experts, leading to computational bottlenecks and increased latency. Therefore, routing policies in MoE–GNN systems should not only optimize prediction performance but also account for system-level metrics such as expert utilization, latency, and communication overhead [195].

From a system perspective, MoE–GNN architectures can be interpreted as distributed, topology-aware processing frameworks. Experts may be deployed across edge nodes, base stations, or centralized controllers, while the routing mechanism determines how graph regions are assigned to different computational units. This raises several practical considerations, including synchronization of expert updates, communication overhead for routing decisions, and adaptation frequency under dynamic network conditions. Consequently, the benefits of MoE–GNN are most pronounced in large-scale, heterogeneous, and highly dynamic environments, whereas in smaller or more stationary networks, the additional architectural complexity may outweigh the gains [198].

Recent studies provide important building blocks toward practical MoE–GNN design in graph-structured and dynamic environments. Complementary works explore additional dimensions of expert design and efficiency. Depth-adaptive MoE distributes expert capacity across network depth to balance expressiveness and computational cost in graph models [199], [200], while spatio-temporal MoE frameworks demonstrate that modular expert design can effectively capture temporal dynamics under non-stationary data regimes [201]. In parallel, movement-aware GNN approaches highlight the importance of selectively modeling spatial dynamics in time-varying networks [202].

Taken together, these developments indicate that the effectiveness of MoE–GNN systems is governed by the joint design of (i) routing granularity (e.g., node-level versus graph-level), (ii) expert specialization (e.g., structural scale, temporal dynamics, or task-specific behavior), and (iii) load balancing mechanisms under dynamic and non-uniform graph distributions. Aligning these components with the structural and temporal characteristics of wireless networks remains a key requirement for achieving scalable, efficient, and robust learning in practical deployments.

### F. World Models–GNN Synergy

World models aim to learn compact, predictive representations of dynamic environments by modeling latent state transitions and enabling forward simulation of system evolution [203]. Formally, a world model encodes observations into a latent state $z_t$ and learns a transition function $p(z_{t+1} \mid z_t, a_t)$, where $a_t$ denotes the control action. This generative state-space formulation enables agents to simulate future trajectories, evaluate alternative decisions, and perform planning without direct interaction with the physical system.

In wireless networks, such predictive capabilities are particularly valuable due to the highly dynamic and stochastic nature of topology, traffic, and interference. Applications such

as digital twin-assisted 6G, mobility-aware resource allocation, and proactive network slicing require models that can anticipate system evolution rather than react to instantaneous observations. Existing work, such as the causal semantic communication framework in [204], demonstrates how world models can capture environment dynamics through imitation learning and causal abstraction. However, these approaches typically lack explicit mechanisms to encode the relational structure inherent in wireless systems.

GNNs provide a natural extension by embedding the underlying network topology into both the representation and transition dynamics of the world model. Specifically, GNN-based encoders can map graph-structured observations (e.g., user–base station associations, interference graphs, or traffic flows) into latent states $z_t$, while graph-based transition models can capture how local interactions propagate through the network over time. This enables the world model to learn structured dynamics such as spatial interference coupling, mobility-driven topology changes, and service dependencies across distributed nodes. As a result, GNN-enhanced world models can generalize across varying network sizes and configurations, which is essential for scalable 6G deployments.

A key advantage of World Models–GNN synergy lies in *generative decision-making*. By leveraging the learned transition model, the system can simulate multiple future trajectories under different action sequences, enabling planning-based control rather than purely reactive policies. For example, a network controller can evaluate the long-term impact of candidate actions (e.g., power allocation, user association, or slice reconfiguration) on performance metrics such as throughput, latency, and energy efficiency. This is particularly relevant in multi-objective and highly coupled wireless systems, where local decisions have global consequences.

Crucially, this paradigm aligns naturally with the control architecture of O-RAN systems. In the near-real-time RAN Intelligent Controller (near-RT RIC), xApps operate on timescales of 10 ms to 1 s and make control decisions based on streaming network telemetry. Integrating a GNN-based world model into this loop enables a predictive control pipeline: (i) telemetry data is encoded into a graph representation of the network state, (ii) the world model predicts future latent states under candidate actions, and (iii) the controller selects actions that optimize long-term objectives based on simulated outcomes. This effectively transforms the RIC from a reactive controller into a model-based decision system capable of anticipatory optimization.

Furthermore, the hierarchical structure of O-RAN allows world models to be deployed at multiple timescales. Near-RT RIC can utilize lightweight GNN-based predictors for fast decision-making, while non-real-time RIC (non-RT RIC) can host more complex generative models (e.g., diffusion or sequence-based world models) for policy training, scenario generation, and digital twin synchronization. This hierarchical integration enables continuous learning and adaptation, where offline-trained world models inform real-time control policies.

Despite these advantages, several challenges remain. Learning accurate and stable graph-based transition models under partial observability and noisy measurements is non-trivial. In addition, the computational overhead of simulating multiple future trajectories must be carefully managed to meet real-time constraints. Finally, ensuring robustness under distribution shifts, such as sudden traffic surges or topology disruptions, requires continual adaptation mechanisms.

Overall, World Models–GNN synergy provides a principled framework for integrating predictive, generative reasoning into wireless network control. By combining graph-structured representation learning with latent state-space modeling, this approach enables scalable, topology-aware, and anticipatory decision-making, aligning closely with the vision of AI-native 6G and intelligent O-RAN architectures.

### G. Toward Unified Intelligence

A key frontier in 6G lies in fusing GNNs with DRL, FL, GNN, MoE, and meta learning into unified frameworks that enable adaptive, decentralized, and cross-layer intelligence. These hybrid models aim to enhance learning efficiency, scalability, and responsiveness in dynamic environments.

Recent efforts exemplify this trend. CFGRL integrates DRL, FL, and GNNs for edge caching under dynamic demands with distributed privacy preservation [205], while FLAG-VNE combines curriculum learning with GNN-based RL for robust virtual network embedding [206]. Multi-layer graph abstractions have also been used to model complex systems like satellite–terrestrial networks for scalable traffic engineering and function placement [207]. Temporal GNNs such as LDP-STG incorporate local differential privacy to track dynamic mobility and traffic patterns [208]. Domain-informed designs like deep graph unfolding embed structural priors to improve convergence and robustness in decoding, power control, and interference mitigation tasks [209], [210], [211], [212]. These developments highlight a shift toward AI-native infrastructures, where GNNs, enhanced by advanced AI, serve as core engines for intelligent, resilient, and self-optimizing 6G networks.

## V. AI–GNNs for Key 6G Technologies

6G is widely envisioned as a shift toward more AI-integrated wireless systems, where learning and inference are increasingly embedded across network architectures, control loops, and protocol layers to enable more autonomous and adaptive operation. To support emerging applications such as holographic telepresence, XR, and Industry 4.0, core technology domains under investigation include intelligent RAN, advanced MIMO with integrated sensing, reconfigurable intelligent surfaces, edge AI, advanced traffic control, heterogeneous coordination, and network DTs. These domains involve large-scale, dynamic, and interdependent relationships among network entities, where graph-based representations, and in particular GNN-based models, offer a natural and effective way to capture relational dependencies and support data-driven optimization.

Alongside these core technologies, next-generation network architectures such as O-RAN provide practical frameworks for realizing AI-driven operation through disaggregation, virtualization, and open interfaces. While O-RAN is not itself

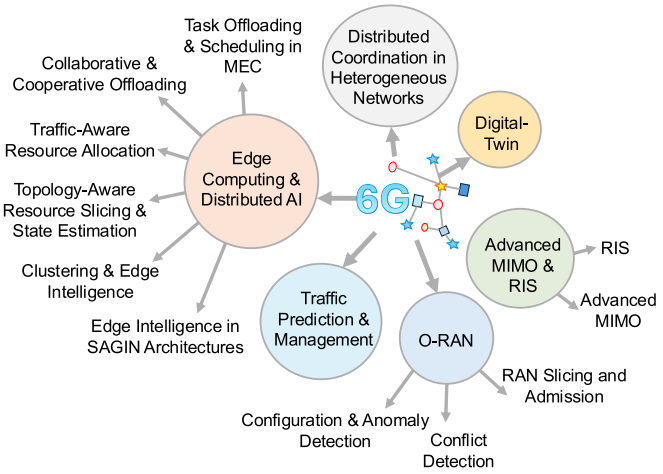


Fig. 5. AI–GNNs Empowering Key 6G Enablers.

a 6G study item, it serves as a representative architectural platform for investigating challenges anticipated in future systems, including distributed intelligence, real-time control, and cross-domain coordination. In parallel, recent 3GPP efforts in Rel-19/20, particularly on AI-native RAN and ISAC, highlight the growing role of learning-based and topology-aware methods in next-generation wireless networks [213].

Within this context, this section examines how GNN-based and GNN-enhanced AI architectures can support selected 6G technology domains and emerging network frameworks. Rather than assuming GNNs as a universal solution, we focus on scenarios where their relational inductive bias provides clear advantages, while also recognizing that alternative modeling approaches may be more suitable in other settings (see Fig. 5).

### A. Next-Generation Network Architectures: O-RAN

The O-RAN architecture exemplifies next-generation wireless network design based on disaggregation, virtualization, and open interfaces. By decomposing the RAN into modular components (e.g., radio, distributed, and centralized units) and virtualizing their functions on commodity hardware, O-RAN enables programmability and flexible deployment [47]. Its RAN Intelligent Controllers (RICs) further support closed-loop optimization and data-driven control.

Although O-RAN is not itself a 6G study item, its architectural characteristics closely align with the requirements envisioned for future 6G systems, such as distributed intelligence, scalable coordination, and low-latency control. In this context, O-RAN provides a practical and extensible platform for deploying GNN-centric AI frameworks. Importantly, the challenges encountered in O-RAN, such as partial observability, coordination across disaggregated components, and real-time constraints, are representative of those faced by next-generation and future 6G network architectures [214]. Accordingly, we use O-RAN as an illustrative case to examine how GNN-based intelligence can enhance control, orchestration, and decision-making in programmable wireless systems.

Although O-RAN is not itself a 6G study item, its architectural characteristics closely align with the requirements envisioned for future 6G systems, such as distributed intelligence, scalable coordination, and low-latency control. In this context, O-RAN provides a practical and extensible platform for deploying AI-driven frameworks, including GNN-based approaches where relational modeling is beneficial. Importantly, the challenges encountered in O-RAN, such as partial observability, coordination across disaggregated components, and real-time constraints, are representative of those faced by next-generation and future 6G network architectures [214]. Accordingly, we use O-RAN as an illustrative case to examine how GNN-based intelligence can enhance control, orchestration, and decision-making in programmable wireless systems.

*1) Ai–GNN for RAN Slicing and Admission:* RAN slicing involves allocating RAN resources to multiple network slices with diverse QoS requirements, and admission control decides whether to accept new slice requests under resource constraints. Conventional approaches (e.g., heuristics or static resource partitioning) are often suboptimal, and early ML solutions using DRL required retraining for each network topology change [215].

GNN-based methods provide an alternative by encoding the RAN as a graph and learning policies that can generalize across different network deployments. For example, a recent work uses a multi-agent DRL framework with GATs to jointly perform slicing and admission decisions [216]. The graph-based agents were topology-independent and operated on arbitrary RAN graphs, achieving improved performance (e.g., up to ~35% higher long-term revenue compared to heuristic slicing and about 20% improvement over non-GNN-DRL baselines). Notably, the GNN-enabled agents maintained strong performance on unseen network scenarios without retraining, indicating improved generalization compared to classical approaches.

*2) Ai–GNN for Conflict Detection:* The deployment of multiple xApps within the Near-RT RIC introduces the risk of policy conflicts, where independently developed control applications operate on overlapping parameters or influence shared KPIs. While conflict management in O-RAN has traditionally relied on rule-based mechanisms or static profiling, recent studies indicate that graph-based learning provides a more expressive and scalable abstraction for capturing the complex and often hidden interdependencies among xApps, control parameters, and KPIs.

Early graph-centric approaches model conflict relationships explicitly but assume that all dependencies are known *a priori*. To overcome this limitation, Zolghadr et al. propose a data-driven GNN-based framework that reconstructs conflict graphs by learning latent relationships between xApps, parameters, and KPIs from operational data using inductive message passing (GraphSAGE) [217]. Their approach enables the identification of direct, indirect, and implicit conflicts without relying on exhaustive sandbox testing, demonstrating the suitability of GNNs for uncovering non-obvious conflict structures in dynamic O-RAN environments.

Building on this line of work, Shami et al. introduce GRACE, a GCN-based framework that explicitly models

xApps, KPIs, and control parameters as a heterogeneous graph to detect policy conflicts in Near-RT RIC deployments [218]. By leveraging relational message passing, GRACE detects direct, indirect, and implicit conflicts and achieves high accuracy on real O-RAN logs, highlighting the generalization advantages of graph-based reasoning over handcrafted rules.

More recently, complementary perspectives have emerged that combine graph representations with explainable learning and causal reasoning. Sharma et al. propose a framework that integrates explainable ML and causal inference to evaluate the impact of conflicting control parameters on KPIs, representing xApp interactions as causal graphs rather than static conflict rules [219]. Although not purely GNN-based, this work reinforces the importance of structured relational modeling for conflict analysis and points toward hybrid graph–causal frameworks for trustworthy conflict management.

Overall, these works highlight the potential of graph-based abstractions for capturing complex interdependencies in O-RAN systems. However, given the evolving nature of this problem space, hybrid approaches incorporating causal inference or explainable models may complement GNN-based methods depending on system requirements.

*3) Ai–GNN for Configuration Recommendation and Anomaly Detection:* Manual configuration and diagnosis of large-scale RANs remain challenging due to the system's high dimensionality, heterogeneity, and evolving dynamics. Traditional rule-based and statistical approaches often lack adaptability to context drift or unforeseen interactions between parameters. While GNN-based solutions for configuration management are still in their early stages, several recent works indicate their potential for enhancing RAN self-optimization, particularly in scenarios where relational dependencies between network entities play a central role.

Piroti et al. [220] proposed a Siamese GNN that learns embeddings from historical configuration and performance data to identify similar "healthy" configurations. By capturing structural relationships across network states, this approach improves KPI compliance and robustness to drift, enabling proactive recommendations and anomaly detection without extensive retraining.

Recent advancements have explored the integration of GNNs with Generative AI for RAN automation. The GNG-OTA framework [221] introduces a GNN-based diagnostic graph combined with generative learning to simulate and evaluate misconfigurations in Open RAN. By modeling KPI propagation across the network, it enables interpretable root-cause analysis and enhances test automation capabilities across multi-vendor environments.

Additionally, a hybrid framework integrating GNNs and large language models (LLMs) has been proposed for configuration verification in 6G-oriented O-RAN systems [222]. In this approach, the GNN component models the structural relationships within the RAN, while the LLM interprets textual configuration intents, bridging the gap between human-defined policies and machine-level automation. This pipeline facilitates natural language-based configuration audits and intelligent policy enforcement, supporting the broader vision of zero-touch network management.

Taken together, these works suggest that GNN-based approaches can serve as a valuable component in scalable and interpretable RAN configuration and diagnostics, particularly when combined with complementary AI techniques. At the same time, their effectiveness depends on the extent to which network behaviors can be meaningfully represented as relational structures, and alternative modeling approaches may be more suitable in scenarios where such structure is less pronounced.

### *B. Advanced MIMO and Reconfigurable Intelligent Surfaces*

Advanced MIMO technologies, including mMIMO, cell-free massive MIMO (CFmMIMO), extremely large aperture arrays, and holographic MIMO, are widely recognized as key enablers of future wireless systems [45], [223]. These architectures aim to exploit spatial degrees of freedom at an unprecedented scale, offering significant gains in spectral efficiency, energy efficiency, and user-centric connectivity. However, their inherently large-scale and distributed nature introduces substantial challenges in channel estimation, beamforming, interference management, and resource allocation. Traditional optimization-based and model-driven solutions often face limitations in scalability, adaptability, and real-time responsiveness. In this context, graph-based representations (in particular, GNN-based learning) provide a promising framework for capturing the structured interactions in MIMO systems and enabling data-driven optimization where such relational structure is prominent.

*1) Advanced MIMO:* Recent research has explored the use of GNNs to address high-dimensional optimization challenges in advanced MIMO architectures such as CF-mMIMO. Unlike conventional optimization or heuristic beamforming strategies, these approaches leverage the graph-structured nature of wireless networks to capture interdependencies among access points (APs), User Equipmentss (UEss), and environmental factors, enabling scalable and low-complexity approximations of otherwise intractable problems.

In the context of ISAC-enabled CF-MIMO, heterogeneous GNN frameworks have been investigated for joint beamforming design. For instance, [224] and [225] employ heterogeneous GNN architectures to jointly optimize communication and sensing performance, demonstrating that such models can balance spectral efficiency and sensing accuracy while reducing reliance on iterative optimization. These works suggest that explicitly modeling structural heterogeneity in the network graph can improve representation capacity and generalization across deployment scenarios.

For resource allocation and power control, several studies indicate that GNNs can approximate solutions to complex combinatorial optimization problems. Parlier et al. [226] formulate optimal linear precoding (OLP) in CF-mMIMO as a node-level prediction task, enabling the GNN to learn dependencies between APs and UEs under practical latency constraints. Similarly, CF-HGNN [227] and distributed approaches such as [228] demonstrate that scalable GNN architectures can support power allocation and user assignment in large-scale CF-mMIMO systems with competitive performance and reduced computational complexity.

Lightweight and hybrid architectures have also been explored to improve deployment feasibility. The multi-head single-body GNN (MHSB-GNN) proposed in [229] incorporates UE-specific updates and prior information to jointly address AP selection and beamforming based on local CSI, offering favorable performance–complexity trade-offs. In addition, model-driven approaches such as [230] combine domain knowledge with data-driven GNN inference, highlighting the potential of hybrid designs in satellite–terrestrial CF-MIMO systems. Collectively, these studies indicate that GNN-based methods can support distributed and adaptive optimization in advanced MIMO settings, particularly when relational interactions dominate system behavior.

*2) Reconfigurable Intelligent Surfaces:* In parallel, GNN-based learning frameworks have been explored for RIS-assisted communications, where the interactions between active transceivers and passive reflecting elements naturally form graph structures. By embedding both geometric and channel-dependent relationships into graph representations, these approaches enable scalable optimization of beamforming, resource allocation, and channel estimation in complex environments.

Heterogeneous and attention-based GNN architectures are commonly adopted in this domain due to their ability to capture diverse and dynamic relationships among RIS elements, users, and base stations. For example, the Meta-path-based Heterogeneous Graph-Transformer Network (MHGphormer) [231] bypasses explicit channel estimation by directly mapping user and environmental features to optimized configurations in RIS-aided THz MU-MIMO systems. Its permutation-invariant design supports varying user densities and sub-band allocations, demonstrating robust performance improvements over classical baselines.

Beyond sum-rate optimization, energy-efficient and distributed solutions have also been investigated. In [232], a GNN-based model learns rate-splitting multiple-access (RSMA) beamforming and power allocation strategies from channel conditions, showing improved energy efficiency and robustness under imperfect CSI compared to DNN baselines. Similarly, [233] integrates heterogeneous GNNs into the joint optimization of active and passive beamforming for STAR-RIS systems, illustrating scalable performance gains under practical constraints.

Unsupervised and attention-based methods further reduce reliance on labeled data and improve adaptability. Works such as [234] employ graph attention mechanisms for precoding design by modeling spatial relationships among RIS elements. In addition, graph-based spatial learning has been applied to channel extrapolation and RIS configuration: [235] uses GCNs with pruning to reduce CSI acquisition overhead, while [236] develops region-specific GNN training for RIS placement and large-scale channel prediction.

GNN-based approaches provide a scalable framework for RIS-assisted systems, particularly when spatial and structural dependencies are prominent. However, their performance depends on the quality of graph representations, and hybrid or model-driven methods may be preferable in scenarios with limited structural regularity.

### C. Edge Computing and Distributed Ai

Edge computing is widely considered a key architectural component of future 6G systems, enabling the decentralization of computation, storage, and intelligence closer to base stations, access points, and end devices [237]. This paradigm supports ultra-low latency, context-aware processing, and real-time analytics, which are important for applications such as XR, industrial automation, autonomous systems, and the tactile internet [238]. However, the highly dynamic, heterogeneous, and resource-constrained nature of edge environments introduces significant challenges in computation offloading, resource orchestration, and mobility management.

Recent advances in machine learning, particularly GNNs, provide a promising framework for addressing these challenges in scenarios where relational structure plays a dominant role. By modeling dependencies across devices, servers, and tasks, GNNs enable topology-aware learning over task graphs, server connectivity, and device interactions. In this section, we examine how GNN-based approaches, often combined with DRL, can support distributed intelligence across MEC infrastructures, while also highlighting the conditions under which such methods are most effective.

*1) Task Offloading and Scheduling in MEC:* MEC brings computational capabilities closer to the UEs, enabling low-latency support for compute-intensive and delay-sensitive IoT applications [46], [239]. In many settings, user tasks can be represented as directed acyclic graphs (DAGs), where nodes denote computational tasks and edges capture dependencies. GNNs are naturally suited to process such structures, enabling data-driven offloading and scheduling strategies that account for both task dependencies and network topology.

Several studies have explored DRL-GNN frameworks for joint task offloading and resource management. For instance, [240] proposes GNNPPO and GNN-MRL, which integrate GNN-based state representations with reinforcement learning to adapt to varying edge conditions, including changing bandwidth, node density, and computational capacity. Similarly, [241] introduces Multi-order Graph Neural Reinforcement Learning (M-GNRL), which captures higher-order spatial correlations through multi-hop aggregation and edge embeddings. These results suggest that GNN-based policies can improve adaptability and generalization in structured MEC environments, particularly when task dependencies and network topology are tightly coupled.

*2) Collaborative and Cooperative Offloading:* In cooperative edge scenarios, where multiple agents share resources and coordinate decisions, GNNs provide a natural mechanism for modeling inter-agent relationships. For example, [242] proposes a Cooperative Offloading for Terminal Clusters (CO-TC) framework for mobile VR systems, combining multi-agent reinforcement learning with graph-based embeddings to jointly optimize task distribution and resource utilization under QoS constraints. Similarly, TransEdge [243] leverages a GNN-based DRL policy to capture spatial dependencies in transportation-edge networks, enabling adaptive node assignment and collaborative offloading.

These approaches demonstrate that GNN-based representations can facilitate distributed decision-making in multi-agent environments. However, their effectiveness depends on the stability of inter-agent relationships and the availability of reliable topology information, which may vary across deployment scenarios.

*3) Traffic-Aware Resource Allocation:* Traffic-aware resource allocation is critical for edge intelligence in vehicular networks and smart city applications. GNN-based temporal models have been explored to capture spatiotemporal traffic dynamics and support proactive resource management. Jiang et al. [244] propose TFFTO, which combines GAT-based traffic prediction with game-theoretic offloading strategies to adapt resource allocation based on predicted demand. Similarly, Song et al. [245] employ a GAT-based temporal graph model enhanced with transformer layers to predict next-day traffic volumes, outperforming traditional time-series approaches.

Ali et al. [246] further introduce DMFGNet, which incorporates a Spatio-Temporal Attention Unit (STAU) to dynamically weight spatial and temporal dependencies. These studies indicate that GNN-based models can enhance traffic-aware decision-making, particularly when spatial correlations and temporal patterns are strongly coupled.

*4) Topology-Aware Resource Slicing and State Estimation:* Highly dynamic MEC environments, such as space-air-ground integrated networks (SAGIN), require adaptive resource slicing under partial observability and time-varying connectivity. In this context, [247] proposes an MPNN-based model to estimate hidden system states, which are then integrated with online ADMM for distributed resource allocation with convergence guarantees. In addition, [248] introduces Distributed Hierarchical Temporal Graph Learning (DHTGL), which combines hierarchical temporal modeling with graph attention to extract spatiotemporal features while reducing communication overhead.

These works suggest that GNN-based state estimation and topology-aware learning can support distributed optimization in dynamic environments. However, their performance is influenced by the accuracy of graph construction and the timeliness of state updates, which remain practical challenges in rapidly changing networks.

*5) Clustering and Edge Intelligence in Vehicular Networks:* GNNs have also been applied to clustering and edge formation in vehicular networks, where topology evolves rapidly due to mobility. For example, [249] proposes a GNN-based clustering framework that captures temporal, spatial, and functional patterns to enable adaptive edge formation. Such approaches highlight the potential of GNNs for maintaining efficient connectivity structures in highly dynamic environments.

*6) Edge Intelligence in SAGIN Architectures:* In SAGIN systems, the integration of satellite, aerial, and terrestrial networks introduces additional heterogeneity and resource constraints. The DRCAM framework [250] leverages knowledge graphs and DRL for computation offloading under multi-dimensional constraints. While not explicitly GNN-based, this line of work indicates that incorporating graph learning models could further enhance relational reasoning and decision-making in such complex environments.

GNN-based methods enable topology-aware and distributed decision-making in edge environments, especially for structured tasks such as task graphs and multi-agent coordination. However, their practicality depends on topology stability, graph construction overhead, and latency constraints, with lightweight or hybrid solutions often better suited for highly dynamic settings.

### D. Traffic Prediction and Management

AI–GNNs provide a flexible framework for modeling network dynamics, which is important for tasks such as traffic prediction, QoS forecasting, encrypted traffic classification, and network anomaly detection. These capabilities can support proactive resource management and improved robustness under diverse 6G network conditions, particularly in scenarios where spatial and temporal dependencies are strongly coupled.

In the domain of QoS prediction, frameworks such as TPP-GNCF integrate GNNs with collaborative filtering and differential privacy to enable privacy-preserving inference in IoT environments [251]. Similarly, QoSGNN models interactions between users and services to provide scalable and fair QoS predictions [252]. Time-aware approaches, such as [253], combine contrastive learning with GCNs to capture user–service–time dynamics, while more advanced models like HCTN [254] incorporate hypergraph convolution and transformers to address temporal sparsity. Additional frameworks, including GACL [255], ChainNet [256], and LightningNet [257], further demonstrate that graph-based learning can improve prediction accuracy and scalability under structured data conditions.

AI–GNNs have also been explored for encrypted traffic classification and network security. TMC-GCN [258] models flow directionality for encrypted traffic classification, while HGNN-ETC [259] captures byte-level temporal relations using high-dimensional graph representations. Similarly, TB-Graph [260] and SAT-Net [261] employ attention-based graph learning to enhance detection of malicious patterns, and GAT-AD [262] applies attention mechanisms for anomaly detection. These results suggest that graph-based representations can improve feature extraction in complex traffic scenarios, although their effectiveness depends on the availability of sufficiently informative graph structures.

For spatio-temporal forecasting and delay prediction, GNN-based models have been combined with attention mechanisms and sequence models to capture multi-scale dependencies. RouteNet-Attention [263] integrates attention into GNNs for delay prediction, while GSSM [264] adopts a graph-based state-space formulation to model base station interactions. Hybrid architectures, such as [265], combine GCNs, transformers, and temporal convolutional networks to improve prediction accuracy in IoT traffic. In vehicular networks, IRNet [266] and FineCDS [267] leverage spatio-temporal graph learning for routing and content delivery, highlighting the role of GNNs in dynamic and mobility-driven environments.

Beyond prediction, AI–GNN approaches have been applied to traffic management and network control. GNNetSlice [268]

enables real-time prediction of latency and reliability metrics for dynamic slice orchestration, while REGNN [269] optimizes packet repetition in URLLC scenarios under interference constraints. In fault detection, BAPT-FL [270] combines GNNs with optimization techniques to localize failures in optical networks. Traffic engineering frameworks such as PathGNN [271] and RDG-TE [272] model routing decisions using graph representations, enabling adaptive and delay-aware optimization in SDN. Similar benefits have been observed in MANETs [273] and emerging vehicular applications such as eGAT [274], as well as in cognitive radio systems using few-shot GNN learning [275].

### *E. Distributed Coordination in Heterogeneous Networks*

AI–GNNs support scalable and adaptive coordination in heterogeneous and ultra-dense wireless environments. Their capacity to represent multi-layered and spatially distributed relationships allows for optimized interference management, energy efficiency, and cross-domain resource coordination.

Recent works demonstrate the applicability of GNN-based methods to power control and interference-aware optimization. For instance, OverGNN [276] employs high-order graph message passing with unsupervised training for power allocation in HUDNs, showing improved energy efficiency under structured interference patterns. Similarly, Edge-GAT [277] models interference using typed edges in heterogeneous graphs, enabling more expressive representations and improved generalization across varying network densities.

Coordination across multiple frequency bands is facilitated through hybrid beamforming schemes. HGNN-HBF [278] models mmWave and sub-6 GHz links using residual attention to guide beamforming. Branch-and-bound acceleration via GNNs [279] significantly reduces the complexity of joint antenna and beam selection while maintaining optimality. RIS-GNN [280] offers a two-phase optimization process tailored for RIS-assisted MISO systems, while HPE-Transformer [281] provides an alternative approach grounded in permutation equivariant (PE) DL for multicast beamforming.

User association and security also benefit from AI–GNN methods. UserBeam-GNN [282] adapts association and beam decisions for users across mmWave and sub-6 GHz systems, improving throughput and system responsiveness. For secure communications, SecGNN [283] applies semantic attention to construct secure beamforming solutions adaptable to varied system constraints.

Load balancing in heterogeneous networks is further addressed by GNN-based models like the LiFi-WiFi architecture [284], which achieves high throughput and reduced latency by adapting to variable access point and user configurations. GFlow [285] merges in-band telemetry with GNN-enhanced DRL to enable fine-grained multipath flow scheduling in MPTCP, improving latency and throughput over conventional methods and showcasing the potential of graph-based AI for intelligent traffic engineering.

### *F. Digital-Twin*

DTs have become central to managing the scale and complexity of emerging 6G networks [286], [287]. Future applications such as extended reality and brain–computer interfaces impose strict latency and reliability requirements that traditional management cannot meet [287]. DTs address this challenge by maintaining real-time virtual replicas of the network, enabling intelligent analysis, prediction, and control through AI and ML techniques.

GNNs naturally complement DTs by modelling relational structures among nodes, links, and traffic flows. Integrating GNNs into DT architectures enhances state representation and predictive accuracy [288], [289], enabling proactive, topology-aware control. For example, a GraphSAGE-based DT improves self-optimization accuracy and convergence speed by encoding node correlations [290], while TwinNet [291] predicts key QoS metrics such as delay and jitter with low error across varying topologies. GNN-powered DTs also strengthen anomaly detection; a 6G edge DT uses flow delay forecasting to identify degradations and trigger GNN-driven reconfiguration [292].

Coupling GNNs with DRL further enables intelligent policy learning within DT environments. In [286], a DT provides predictive feedback for multi-agent DRL, jointly optimizing sensing accuracy and offloading energy. In MEC systems, DT-driven DRL frameworks improve task offloading efficiency and revenue [293]. DTs have also been applied to network slicing [294] and dynamic service placement [295], offering faster adaptation and improved resource allocation.

To enhance generalization across dynamic 6G scenarios, Transfer Learning (TL) and Multi-task Learning (MTL) have been incorporated into DT designs. TL accelerates service migration by reusing prior DRL policies rather than retraining [295], while MTL-based DTs support multi-task driver behavior analysis, improving overall accuracy [296]. These strategies reduce overhead and speed up intelligent twin deployment.

Overall, GNNs and advanced AI techniques greatly amplify the capabilities of next-generation DTs for 6G. They enable high-fidelity modeling of network behavior, proactive detection of issues, and autonomous optimization, all in a topology-aware manner that traditional flat models cannot achieve. Importantly, these AI-driven twins enhance network resilience (through self-healing and anomaly prediction [292] and efficiency (through knowledge reuse and faster decision-making [297]. As 6G networks move towards intent-based and self-X (self-configuring, self-optimizing, self-repairing) operation, graph-based DTs are poised to become indispensable tools for network operators, bridging the gap between complex physical networks and intelligent automated management systems.

Table IV summarizes representative works across these domains by linking application tasks with their corresponding AI–GNN architectures and practical roles.

## VI. Challenges in Integrating AI–GNNs Into 6G

The fusion of GNNs and advanced AI paradigms with 6G wireless systems represents a transformative shift in network intelligence, moving from peripheral optimization to deeply embedded cognition. Unlike prior generations that relied on largely modular control logic, 6G envisions intelligence as a native, distributed, and context-aware capability spanning

TABLE IV
COMPREHENSIVE SUMMARY OF AI–GNN ARCHITECTURES ACROSS KEY 6G TECHNOLOGY DOMAINS

| Domain | Task | AI–GNN Paradigm | Key Idea | Practical Advantage | Ref. |
|---|---|---|---|---|---|
| **O-RAN / Intelligent RAN** | | | | | |
| RAN slicing | Slicing + admission | GNN + DRL | Dynamic orchestration over time-varying graphs | Improved slice embedding accuracy | [154] |
| RAN slicing | Slicing + admission | GNN + DRL | Graph-aware slice embedding over heterogeneous RAN graph | Improved service quality and embedding accuracy | [155] |
| RAN slicing | Slicing + admission | GAT + MARL | Topology-aware policy learning over RAN graphs | Generalization across deployments; no retraining | [216] |
| Conflict Mgmt | Causal conflict analysis | Graph + causal ML | Model KPI interactions via causal graphs | Interpretability + root-cause insight | [219] |
| Config Mgmt | Config recommendation | Siamese GNN | Learn similarity between "healthy" configurations | Drift robustness, proactive optimization | [220] |
| Config Mgmt | Misconfiguration diagnosis | GNN + Generative AI | Simulate KPI propagation via diagnostic graph | Explainable root-cause analysis | [221] |
| Config Mgmt | Policy verification | GNN + LLM | Combine structural graph + textual intent | Natural language auditing, zero-touch management | [222] |
| **Advanced MIMO / RIS** | | | | | |
| CF-MIMO | Power allocation | HGNN / distributed GNN | Scalable resource allocation over AP-UE graph | Reduced complexity, distributed operation | [227] |
| RIS | Beamforming (THz MU-MIMO) | Heterogeneous attention-based GNN | Direct mapping without CSI estimation | Robust to user density variations | [231] |
| STAR-RIS | Precoding | Unsupervised and attention-based | Optimize STAR-RIS configurations | Scalable optimization | [234] |
| **Edge Computing / MEC** | | | | | |
| MEC | Task offloading | GNN + DRL | Task dependency graph modeling for energy-aware offloading | Energy-efficient multi-user offloading | [158] |
| MEC | Task offloading | GNN + DRL | Graph convolutional RL for UAV aerial edge computing | Joint trajectory and task optimization | [159] |
| MEC | Task offloading | GNN + DRL | Dependent task offloading using GNN embeddings | Improved service utility | [161] |
| MEC | Task offloading | GNN + PPO / MRL | Graph-based state representation for DRL | Adaptation to dynamic conditions | [240] |
| MEC | Offloading (hierarchical) | Multi-order GNN + RL | Multi-hop feature aggregation | Capture higher-order dependencies | [241] |
| MEC | Cooperative offloading | GNN + Multi-agent RL | Model inter-agent relations | Energy-efficient coordination | [242] |
| MEC | Cooperative offloading | GAT + Multi-agent RL | Content caching in vehicular edge using graph attention RL | Reduced caching latency | [157] |
| MEC | Edge caching | GNN+DRL+FL | Unified distributed privacy-preserving edge caching | Dynamic demands with privacy preservation | [205] |
| MEC | Transportation-edge | GNN + DRL | Spatial dependency modeling | Improved task assignment | [243] |
| MEC | Traffic-aware offloading | GAT + game theory | Predict traffic $\rightarrow$ optimize offloading | Proactive resource allocation | [244] |
| MEC | Traffic prediction | GAT + Transformer | Spatio-temporal modeling | Improved forecasting accuracy | [245] |
| MEC | Resource prediction | ST-GNN (DMFGNet) | Spatio-temporal attention | Real-time + offline prediction | [246] |
| MEC | State estimation | MPNN + ADMM | Hidden state inference + optimization | Convergence guarantees | [247] |
| MEC | Distributed learning | GAT + temporal model | Reduce communication overhead | Efficient distributed inference | [248] |
| SAGIN | Offloading | GNN + DRL +meta learning | Multi-dimensional decision modeling | Fast adaptation under changing conditions | [186] |
| SAGIN | Offloading | Knowledge graph + DRL | Multi-dimensional decision modeling | Extendable to GNN frameworks | [250] |
| **Traffic Prediction and Management** | | | | | |
| QoS | QoS prediction | GNN + CF + DP | Privacy-preserving inference | Scalable + secure prediction | [251] |
| QoS | Temporal QoS | GCN + contrastive | Capture time dependencies | Better temporal modeling | [253] |
| QoS | Sparse QoS | Graph + Transformer | Handle data sparsity | Improved robustness | [254] |
| Security | Byte-level modeling | HGNN | Capture packet structure | Fine-grained classification | [259] |
| Forecasting | Delay prediction | GNN + Attention | Model network paths | Accurate delay estimation | [263] |
| Traffic Eng | Routing | GNN + DRL | Path-link modeling | Adaptive routing decisions | [272] |
| HUDN | Power allocation | GNN (unsupervised) | High-order message passing | Energy efficiency | [276] |
| Beamforming | Hybrid BF | HGNN | Multi-band modeling | Efficient coordination | [278] |
| Security | Secure BF | GNN + attention | Semantic-aware design | Robust security | [283] |
| Traffic Eng | Flow scheduling | GNN + DRL | In-band telemetry integration | Low-latency optimization | [285] |
| Traffic Eng | UAV Scheduling | GNN + DRL | UAV dynamic scheduling | Improved scheduling in mobile edge | [165] |

TABLE IV
*(Continued.)* COMPREHENSIVE SUMMARY OF AI–GNN ARCHITECTURES ACROSS KEY 6G TECHNOLOGY DOMAINS

| | | | | | |
|---|---|---|---|---|---|
| **Digital Twin** | | | | | |
| DT | Control | GNN + DRL | Policy learning in twin | Joint optimization | [286] |
| DT | Transfer learning | GNN + TL | Knowledge reuse | Faster adaptation | [295] |
| DT | Task offloading | GNN + DRL | Digital twin-driven intelligent task offloading in MEC | Improved offloading efficiency and revenue | [293] |
| DT | Network slicing | GNN + DRL | Digital twin-assisted resource allocation for slicing | Faster adaptation and improved resource allocation | [294] |
| DT | Self-healing | GNN + DRL | Flow delay forecasting for anomaly-triggered reconfiguration | Improved network reliability and resilience | [292] |

the entire protocol stack. In this paradigm, GNNs provide a compelling inductive basis for modeling the dynamic and irregular structures inherent to wireless networks, including user associations, interference patterns, mobility graphs, and service topologies. When hybridized with complementary paradigms such as DRL, FL, meta-learning, generative AI, and neurosymbolic learning, AI–GNN systems promise scalable, adaptive, and topology-aware intelligence for both edge and core network decision-making.

Despite this promise, integrating AI–GNNs into real-world 6G deployments poses numerous challenges, not just in implementation but in foundational research. These include constraints on scalability, real-time adaptation, generalization to unseen or evolving topologies, and interpretability in safety-critical and policy-bound environments. Additionally, architectural hurdles, such as heterogeneity of devices and interfaces, symbolic constraint modeling, and efficient distributed execution, remain unresolved. This section critically examines these challenges and outlines the directions in which further work is needed.

### A. Scalability and Real-Time Constraints

A key barrier to the practical deployment of AI–GNN frameworks in 6G systems is achieving real-time inference and learning at scale in ultra-dense environments. As future 6G networks may interconnect millions of heterogeneous nodes, ranging from IoT sensors and user devices to vehicles and network elements, GNN-based models face exponential growth in graph size, resulting in substantial memory and computational demands [167], [298].

In deep GNN architectures, graph convolutions require aggregating information from increasingly large neighborhoods, causing the computational complexity of both training and inference to grow rapidly with network density. This aggregation overhead can become prohibitive, particularly in dynamic scenarios, risking latency violations in time-critical use cases such as URLLC, real-time routing, resource allocation, and dynamic beam coordination [148], [299].

To mitigate these challenges, a range of lightweight approximations, including neighborhood sampling, clustering-based aggregation, hierarchical graph partitioning, and edge pruning, have been proposed to reduce computational load and memory footprint [300]. However, these techniques typically introduce accuracy–complexity trade-offs and may degrade performance under highly dynamic network conditions.

More advanced solutions are therefore emerging, including hierarchical and distributed GNN processing across cloud–edge infrastructures, where partial inference is performed locally and refined centrally, as well as streaming and incremental GNNs that update node embeddings on-the-fly without recomputing the full graph [301], [302]. In parallel, progress in specialized hardware accelerators and GNN-aware processing units is enabling more efficient execution of large-scale graph workloads [303]. Ultimately, achieving millisecond-scale AI–GNN responsiveness in rapidly evolving 6G environments will likely require a holistic synthesis of algorithmic sparsification, adaptive depth and receptive-field control, edge-native model distillation, and hardware–algorithm co-design.

### B. Generalization and Robustness

Closely tied to scalability is the challenge of generalization; an AI–GNN model trained in one network scenario must adapt to new, larger, or differently structured networks. Studies have shown that vanilla GNNs often suffer poor accuracy when applied to topologies larger or more complex than those seen in training [304], [305]. In other words, a model learned on a small-scale network may not scale out-of-the-box to a city-wide 6G mesh with far more nodes or new traffic patterns [306].

Addressing this calls for more flexible GNN architectures and training paradigms [307]. Techniques such as meta-learning, TL, or graph augmentation using fundamental networking domain knowledge have shown promise in improving GNN generalization [307], [308]. For instance, incorporating network routing principles into GNN design has significantly improved delay prediction accuracy on larger, unseen networks [309].

Moreover, 6G networks will be highly dynamic, with nodes joining and leaving, channels fluctuating, and traffic demands evolving, so GNN models must continually adapt online [310]. Current GNN frameworks struggle with such streaming adaptation, often requiring costly re-training when the graph changes.

Research is turning toward dynamic GNNs that update their parameters incrementally as the topology evolves. This is still nascent, but early examples in THz vehicular networks show the potential: a dynamic GNN that adjusts its message-passing strategy per topology was able to serve 17% more vehicles in a joint communication-sensing task compared to a static

model [163]. GNNs' ability to handle spurious subgraphs improves the robustness and trustworthiness of GNNs in wireless communication [311]. Such adaptability will be essential for GNNs to remain robust in the face of 6G's non-stationary conditions.

### C. Interpretability and Trustworthiness

The closed-box nature of deep GNNs presents a significant barrier to adoption in mission-critical 6G systems. Network operators must trust AI-driven decisions, such as user prioritisation, handover triggers, or slice reconfiguration, yet AI–GNN models typically compress inputs into latent representations that obscure the underlying rationale [312]. When an AI–GNN-based controller misallocates resources, diagnosing the cause is often difficult, hindering both operational safety and regulatory compliance.

Early efforts to enhance transparency include attention-based GNN variants that highlight influential neighbours or links [313], revealing, for example, which user's channel quality dominated a handover decision [314]. Post-hoc methods such as GNNExplainer, PGExplainer, GraphLIME, and GraphMask provide node-, feature-, or subgraph-level attributions, while tools such as Generative Adversarial Network (GAN) Explainer [315] and SHAP-based approaches [316] offer more general perturbation-driven explanations. These methods increase interpretability but remain insufficient for fully accountable AI-driven networking.

Recent work explores combining explainability with structured learning mechanisms. A notable example is TANGO, a symbolic GRL framework that integrates Bayesian GNN explanation with logic-based reward shaping, improving both transparency and convergence for gNodeB-level resource allocation [314]. Such hybrid approaches demonstrate that interpretable reasoning can coexist with high-performance GNN-based control.

Additional evidence for the value of explainable learning in wireless systems comes from hypergraph-inspired models. In UAV-assisted IIoT beamforming, Zhao et al. propose a hypergraph-augmented explainable deep learning framework that explicitly captures heterogeneous interference patterns and preserves structural information lost in conventional graph models. Their results show improved scalability, stability under varying power constraints, and more interpretable resource-allocation behaviour [317]. This reinforces the broader trend that explainability, particularly when combined with richer relational representations such as hypergraphs, is becoming essential for trustworthy 6G AI.

### D. Dynamic Graphs and Mobility

6G will operate over highly dynamic and partially observable topologies, including satellite constellations, UAV fleets, and vehicular swarms. AI–GNN systems must function under continual structural evolution, with intermittent link failures, sparse feedback, and partial observability [318]. While dynamic GNNs like EvolveGCN and DySAT can handle evolving connectivity, most models still assume periodic, synchronized updates and complete topology snapshots. Realistic 6G deployments demand asynchronous, event-driven, and lightweight update mechanisms that support fast reactivity and fault tolerance. Emerging models combine recurrent and attention-based structures for this purpose, but more research is needed on edge-deployable dynamic GNNs that function under lossy, decentralized conditions.

### E. Technical and Architectural Limitations

Heterogeneity across nodes, protocols, and hardware platforms is intrinsic to future 6G networks and poses a fundamental challenge for current GNNs. Wireless systems are expected to interconnect diverse entities, including mobile devices, base stations, aerial platforms, satellites, and large-scale IoT deployments, forming highly heterogeneous and multi-relational interaction graphs. Most existing GNNs, however, assume homogeneous node types and uniform connectivity patterns, limiting their ability to faithfully model such complex environments. Recent advances in heterogeneous GNNs aim to address this limitation by supporting multi-type nodes and relations (e.g., terrestrial–aerial–satellite graphs), yet achieving expressive modeling while maintaining scalable and efficient inference remains an open problem [319].

Beyond structural heterogeneity, incorporating domain knowledge and hard system constraints into AI–GNN models remains challenging. Wireless networks operate under strict physical-layer, protocol, and regulatory constraints, including power budgets, spectrum masks, latency bounds, and QoS guarantees. Many deep learning–based approaches learn unconstrained solutions that must subsequently be clamped or corrected to ensure feasibility, undermining reliability and interpretability. For instance, GNNs trained for power control or scheduling may output decisions that violate transmitter power limits or timing constraints unless manually adjusted [320].

While several studies attempt to encode constraints via tailored loss functions or architectural modifications, such approaches are often brittle and scenario-specific [321]. A promising research direction involves neurosymbolic integration, in which differentiable GNN modules are coupled with symbolic reasoning, logical rules, or optimization-based constraints to enforce feasibility by design [322].

From an architectural and deployment perspective, integrating AI–GNN models into real-world network control loops introduces additional challenges. In disaggregated and softwarized RAN architectures, AI agents are envisioned to operate within near-real-time and non-real-time controllers (xApps and rApps), interacting with the network via standardized interfaces.

However, practical deployment is hindered by heterogeneous hardware capabilities across radio units, edge nodes, and centralized processing units, as well as by interface latency, limited observability, and the absence of mature runtime standards for AI execution. Efforts toward standardized abstraction layers (e.g., O-RAN O1/E2/A1 interfaces) and portable execution frameworks (e.g., ONNX-based models and containerized ML runtimes) provide important enablers, but bridging the

gap between theoretical AI–GNN designs and robust, real-time deployment remains an open challenge. Addressing this gap will require advances in abstraction-robust model design, cross-layer co-optimization, and hardware-aware GNN compilation to ensure reliability, interoperability, and performance at scale in production 6G systems.

In summary, the adoption of AI–GNNs in 6G networks hinges on addressing a multi-dimensional challenge space. From algorithmic scalability and structural generalization to real-time inference, explainability, and symbolic integration, hybrid GNN-based solutions must evolve toward robustness, transparency, and deployability to fulfill the vision of AI-native, graph-empowered wireless systems.

## VII. Emerging Applications and Visionary Uses of AI–GNNs

Despite the challenges, GNN-Based AI is unlocking visionary 6G applications that could radically transform wireless services:

### A. Semantic and Goal-Oriented Communications

6G aspires to move beyond bit-pipe communications toward semantic communication, where networks convey meaning and goals rather than raw data. AI–GNNs are proving instrumental in this shift. One intriguing approach uses knowledge graphs and GNN encoders to represent information in terms of semantic relationships, enabling ultra-efficient compression and interpretation of messages [121]. For example, instead of transmitting a verbose sensor dataset, a 6G device could send a compact graph of key concepts (nodes) and their relationships (edges) essentially a semantic snapshot that an AI at the receiver can understand and act upon. Recent work achieved high compression by cascading large language models with GNN-based semantic encoders, transmitting only node embeddings and reconstructing the full knowledge graph at the receiver.

This kind of graph-based semantic encoding dramatically reduces transmitted bits while ensuring the receiver grasps the intended meaning. In another visionary leap, researchers are combining GNNs with Generative AI to enable immersive multimodal communications for the metaverse. In a cross-modal semantic communication system, a GNN first distills multimodal inputs (e.g. video, audio, haptic data) into a graph of salient features, which is then compressed and sent wirelessly [187].

On the receiving end, generative models use this graph to reconstruct rich media (like 3D holographic scenes or AR/VR environments).Early prototypes show substantial potential, such GNN-assisted semantic pipelines enhanced user interaction in a 6G metaverse scenario, while keeping bandwidth use efficient. In summary, AI–GNNs could empower 6G networks to communicate what things mean or intend, rather than just exchanging bits, laying the groundwork for goal-driven services (e.g. instructions to autonomous drones or summaries for AI assistants) with unprecedented communication efficiency.

### B. Semantic Awareness in ISAC

As 6G evolves toward becoming an intelligent, context-aware infrastructure, semantic awareness, the ability of networks to understand and reason about the meaning behind data, signals, and user behaviour, has become a central theme. A key enabler of this shift is ISAC. This paradigm supports use cases like vehicles that both communicate and perceive their surroundings, or base stations that detect user presence while delivering connectivity, executing both functions semantically, in a coordinated and context-informed way.

GNNs have emerged as a powerful tool for enabling such semantic understanding in ISAC systems. ISAC tasks often depend on relational reasoning, deciding which nodes (e.g., base stations, vehicles, sensors) should take on sensing roles, communication roles, or both, based on dynamic context. GNNs are naturally suited for this, as they model structured interactions in a graph where nodes and edges carry semantic meaning. For instance, in a recent study on terahertz vehicular networks, a GNN-based approach was used to semantically allocate vehicle-to-infrastructure links for both sensing and data services [125]. Vehicles and roadside units were modeled as graph nodes with weighted links reflecting service potential. The GNN dynamically learned which vehicles should receive communication, sensing, or hybrid services, outperforming static heuristics and adapting well to fast-changing road environments. More broadly, GNNs can capture multi-modal semantic context, like identifying clusters of sensors whose data is redundant, or routing traffic through nodes with better situational awareness due to co-located sensors. This semantic fusion of sensing and communication allows the network to act with greater intent and efficiency.

In future 6G smart city deployments, for example, a GNN could analyse a semantic graph of street cameras, mobile users, and environmental sensors to jointly make decisions about handoffs, radar angles, and resource allocation. The GNN's holistic view of the network as an interactive, context-rich graph enables intelligent coordination across layers, bridging perception and connectivity in a way that static systems cannot. Ultimately, semantic awareness powered by GNNs will be central to unlocking the full potential of ISAC, transforming networks from passive data pipes into cognitively aware systems that adapt, reason, and optimize based on the meaning of the information they carry and sense.

### C. Autonomous Network Operation and Self-Optimization

Perhaps the most transformative vision for 6G is that of a fully autonomous network that can monitor itself, learn, and optimize without human intervention. AI-powered GNNs are poised to be central in this evolution, essentially serving as the "neural brains" of self-driving networks. One concrete application is in network self-healing: detecting faults or performance drops and dynamically reconfiguring the network to compensate.

DT technology is converging with GNNs to facilitate this. A DT creates a live digital replica of the network (e.g. a graph model of all nodes, links, and their states), which a GNN can continuously analyse to spot anomalies and recommend

adjustments [323]. For instance, if a set of 6G edge servers becomes overloaded or a link fails, the GNN (running on the twin) could predict how to reroute traffic or spin up backup nodes, effectively healing the service in real-time. Early experiments in a 6G edge testbed demonstrated a GNN-driven digital twin that could trigger service reconfiguration upon failures, significantly improving network reliability [44].

Beyond fault management, AI–GNNs can handle resource optimization tasks autonomously, such as tuning power allocations, spectrum assignments, or antenna configurations on the fly as conditions change. In O-RAN architectures, for example, researchers have used GNN-based xApps to optimize user association and load balancing, treating the radio access network as a graph of connections that can be intelligently rewired in real time. The result is a network that learns from experience (past traffic patterns, mobility events) and preemptively adjusts to meet performance goals (throughput, latency, reliability) without waiting for human planners.

Another emerging application is network slicing and multi-domain orchestration: 6G will span terrestrial, aerial (drones, HAPS), and satellite segments, and GNNs are being investigated to coordinate end-to-end slices across these domains. By modeling inter-domain dependencies as a graph, a GNN agent could allocate resources across a whole 6G ecosystem to meet a slice's Service Level Agreement (SLA). While these autonomous networking applications are still in trial stages, they epitomize the 6G vision of a self-driving network. GNN-Centric AI Architectures essentially imbues the network with a form of collective intelligence, enabling it to perceive context (through distributed data), reason over its graph-structured state, and act in a closed control loop. If realized, this would fundamentally reshape how future wireless systems are operated, shifting from static, human-defined configurations to fluid, AI-guided orchestration that continuously evolves.

### D. Case Study: GraphNet, FGraphNet

This subsection presents an illustrative case study to demonstrate how hybrid GNN-centric architectures can be applied to wireless channel estimation. Rather than introducing new experimental results, this discussion highlights architectural insights and draws on empirical findings reported in prior works [324], [325]. It complements the conceptual discussion in Subsection IV-B by grounding the synergy between GNNs and distributed learning in concrete system designs. Early deep learning-based channel estimation approaches predominantly relied on CNNs, leveraging the regular time–frequency grid structure of resource elements. While effective under quasi-stationary channel conditions, these methods struggle to generalize in highly dynamic environments due to their reliance on fixed local receptive fields and dense pilot patterns. To address these limitations, GraphNet reformulates channel estimation as a relational inference problem using graph-based representations. In this framework, resource elements are modeled as nodes, and dependencies between pilot and data symbols are captured as edges, enabling flexible information propagation across the graph. This shift replaces rigid Euclidean inductive biases with topology-aware representations, thereby improving robustness under irregular

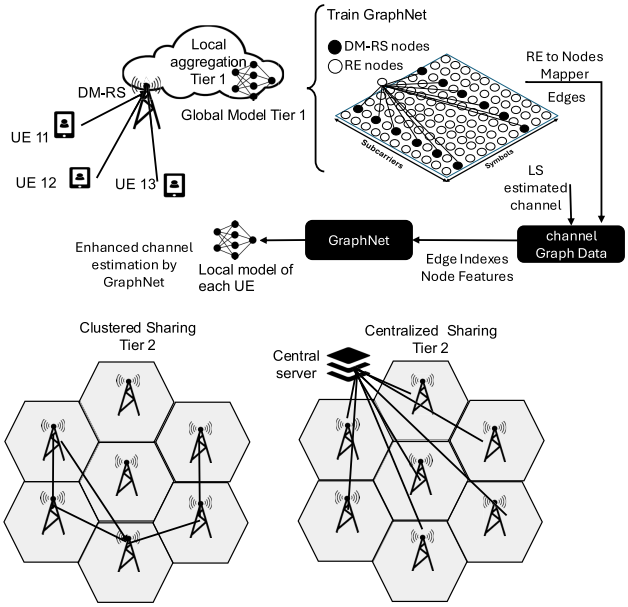


Fig. 6. GraphNet and Two-Tier Federated GraphNet (FGraphNet) for Scalable Channel Estimation in 5G NR. Local channel data are transformed into graph representations and used to train GNN-based channel estimators at each BS (Tier-1). Model updates are then shared across neighboring BSs within a cluster or centrally (Tier-2), enabling collaborative and adaptive channel estimation under heterogeneous network conditions.

pilot distributions and rapidly varying channel conditions. Simulation results reported in [324] show that GraphNet achieves improved BLER and MSE performance compared to CNN-based approaches such as ChannelNet [326] in dynamic environments, while maintaining a model size over 250× smaller, making it suitable for real-time edge deployment.

Beyond centralized learning, this case study further illustrates how GNNs synergize with FL to address scalability, heterogeneity, and privacy constraints in wireless networks. Building on the graph-based channel estimation formulation, FGraphNet introduces a decentralized, two-tier learning framework in which GNN models are trained locally at BSs and collaboratively aggregated without sharing raw channel data. This approach enables adaptive learning across diverse channel conditions and user mobility profiles while preserving data locality. Simulation results in [325] demonstrate that this federated graph-based framework reduces MSE, accelerates convergence, and lowers local training time through collaborative knowledge sharing across base stations. These results highlight that the benefits of GNN-based channel estimation extend beyond model design to system-level efficiency and scalability when integrated with decentralized learning paradigms. Overall, this case study highlights a broader insight for AI-native wireless systems: by combining graph-based relational modeling with distributed learning paradigms, GNN-centric architectures enable robust, scalable, and cooperative network intelligence that extends beyond isolated performance gains to system-level adaptability.

## VIII. Research Gaps and Future Directions

To unlock the full potential of AI–GNNs in 6G, further research must tackle several open issues and push the boundaries of current technology:

### A. Energy Efficiency of Ai–GNN Workloads and the Imperative of Sustainability

The deployment of GNN-Centric AI Architectures at scale introduces serious concerns about energy consumption, raising critical questions about the sustainability of future intelligent networks. Paradoxically, while GNNs can help optimize network energy usage, such as by intelligently switching off underutilized base stations, they themselves require significant computational power. Training and inference of deep GNNs on large-scale 6G graphs can consume substantial energy if not carefully designed and managed [327]. Additionally, energy-aware training, where energy consumption is embedded into the loss function, could steer AI models toward not only high accuracy but also low-power operation. As 6G networks increasingly embrace sustainability as a first-class goal [328], aligning AI–GNN innovation with environmentally responsible design principles will be a defining challenge and opportunity for the next generation of network intelligence.

### B. Privacy-Preserving and Distributed Learning

The ubiquity of 6G networks is expected to generate massive volumes of data, ranging from fine-grained user context and mobility traces to dense sensor streams and network telemetry. While such data are invaluable for training powerful GNNs, centralizing them at the cloud raises serious privacy, security, and bandwidth concerns. In many emerging applications, such as autonomous driving, industrial IoT, and personalized services, sharing raw data is either impractical or fundamentally incompatible with regulatory and ethical constraints.

To address these challenges, distributed GNN learning has emerged as a key research direction, enabling collaborative model training across edge devices, access points, or base stations without exposing sensitive data. FL fits naturally into this paradigm, where each participant trains a local GNN on its regional or task-specific graph and shares only model updates with an aggregator [329]. Recent frameworks demonstrate the feasibility of federated GNNs by exchanging gradients or embeddings rather than raw graph data, achieving strong performance while preserving data locality [330].

Despite this promise, federated and decentralized GNN training introduces challenges that are fundamentally more complex than those encountered in conventional federated learning. Graph-structured data across network nodes are often highly heterogeneous, non-IID, and dynamically evolving, reflecting differences in topology, user behavior, and traffic patterns. Such heterogeneity can significantly degrade model convergence, generalization, and personalization, particularly in mobile and large-scale 6G environments [331], [332]. Designing aggregation mechanisms and personalization strategies that explicitly account for graph diversity remains an open research problem.

Communication and resource efficiency constitute another critical bottleneck. Synchronizing high-dimensional GNN parameters or node embeddings over large and dynamic graphs can incur substantial communication and computational overhead, straining bandwidth-limited wireless links and resource-constrained edge devices. While techniques such as sparse updates, asynchronous aggregation, model compression, and split learning have been proposed to alleviate these costs, they introduce tradeoffs between learning accuracy, convergence speed, and system overhead [333]. Developing communication-aware and system-aligned training protocols remains essential for scalable deployment.

Privacy concerns extend beyond training to inference and model usage. Even after training, GNN-based decision-making can leak sensitive information through model outputs or intermediate representations. To mitigate this risk, techniques such as differential privacy, homomorphic encryption, secure multi-party computation, and trusted execution environments are increasingly explored [334]. For example, recent work in vehicular networks has demonstrated privacy-preserving attention-based GNNs for task offloading, achieving low-latency decision-making while protecting vehicle-level data [335]. However, such protections often come at the cost of increased latency or reduced accuracy, highlighting unresolved privacy–utility tradeoffs.

Security and robustness further complicate distributed GNN learning. Graph topology itself can be a target of adversarial manipulation, including poisoning, Sybil attacks, and structural perturbations during federated training [336]. Existing defenses remain limited, motivating future research into adversarially robust federated GNNs, anomaly-resilient aggregation, and trust-aware coordination mechanisms. The integration of blockchain or distributed ledger technologies has been proposed to enhance auditability and trust, but their scalability and overhead in 6G contexts remain largely unexplored [337].

### C. Neuro-Symbolic Integration for GNNs in 6G

To overcome the limitations of purely data-driven models, neuro-symbolic AI, the fusion of neural networks with symbolic reasoning and domain knowledge, is gaining traction. In the 6G context, this means embedding protocols, policies, and constraints directly into GNN-based models to ensure compliance and improve interpretability [322]. For example, symbolic modules can enforce hard rules, such as ensuring GNN-predicted power levels stay within regulatory bounds or that routing decisions align with reliability requirements. Knowledge graphs can also be integrated to represent network functions, spectrum policies, or service hierarchies, allowing GNNs to reason over structured domain knowledge.

Early work in semantic communication has demonstrated this concept, using knowledge graphs with GNNs and language models to guide encoding decisions. This points toward a broader vision: wireless controllers that not only learn from data but also consult symbolic knowledge to align with operator-defined goals.

The main challenge lies in designing architectures that blend learning and reasoning. One promising direction is a hybrid control system, where a GNN proposes actions and a symbolic engine validates or adjusts them, balancing optimality with rule compliance. Such integration also improves explainability, enabling models to justify decisions using human-readable logic (e.g., “Device X was deprioritized due to policy Y”) rather than opaque activations. Though still

in early stages, neuro-symbolic GNNs align well with 6G's vision of intelligent, transparent, and controllable networks, and will require close collaboration between AI researchers and wireless experts to operationalize.

### D. Real-World Deployment Challenges

While much of the existing literature focuses on algorithmic performance, deploying GNN-based solutions in operational 5G and emerging 6G systems requires substantial system-level redesign and introduces several practical challenges for mobile network operators (MNOs).

*1) Integration Into Existing 5G/O-RAN Control Loops:* A key challenge lies in embedding GNN-based intelligence into existing network architectures. For example, in O-RAN systems, control logic is executed through Near-RT and Non-RT RICs via xApps and rApps. Integrating a GNN-based scheduler or slicing controller requires redesigning these control loops to (i) construct graph representations from streaming telemetry (e.g., KPM, RC data), (ii) ensure consistent state synchronization across distributed units (O-DU, O-CU), and (iii) meet strict latency budgets (typically 10–100 ms in Near-RT RIC). Unlike traditional rule-based xApps, GNN-based solutions require continuous graph updates and inference pipelines, which introduces additional system complexity and overhead.

*2) Data Collection and Graph Construction:* Effective GNN deployment depends critically on the availability of structured and high-quality graph data. In practice, network data are distributed across multiple sources (e.g., base stations, edge nodes, core network functions) and often exhibit inconsistencies, missing values, or delayed reporting. Constructing reliable graphs from such telemetry—defining nodes, edges, and features in a consistent manner—remains a non-trivial task. Moreover, real-time graph construction can introduce significant processing overhead, particularly in large-scale networks with thousands of nodes and rapidly changing topology.

*3) Model Lifecycle Management:* Unlike static optimization algorithms, GNN-based models require continuous lifecycle management, including training, validation, deployment, monitoring, and updating. In operational networks, traffic patterns, user behavior, and topology evolve over time, leading to model drift. MNOs must therefore implement mechanisms for online adaptation, periodic retraining, and rollback in case of performance degradation. This raises challenges related to data versioning, model validation under safety constraints, and integration with existing network management and orchestration (MANO) frameworks. Ensuring reliability and stability of AI-driven decisions remains a critical concern, particularly for safety-critical services.

*4) Edge Deployment and Hardware Acceleration:* Many GNN-based applications, such as scheduling, resource allocation, or anomaly detection, require inference at the edge under strict latency and energy constraints. However, GNNs are inherently memory- and communication-intensive due to neighborhood aggregation and irregular data access patterns. Deploying them on edge hardware (e.g., baseband units or edge servers) without acceleration can lead to unacceptable latency. While emerging solutions such as GPU-based frameworks, FPGA implementations, and specialized accelerators (e.g., GraphCore IPU) offer potential improvements, their integration into telecom infrastructure is still limited. Furthermore, hardware-aware GNN design, including model pruning, quantization, and sparse computation, remains an active research area.

*5) Operational and Economic Considerations:* From an MNO perspective, adopting GNN-based solutions must be justified in terms of operational complexity, cost, and reliability. Introducing AI-driven control requires additional infrastructure for data pipelines, model management, and monitoring, as well as new expertise in AI engineering. Moreover, explainability and trustworthiness are essential for deployment in carrier-grade networks, where decisions must be auditable and compliant with regulatory requirements. These factors may limit the immediate adoption of complex GNN models in favor of simpler or hybrid solutions.

In summary, bridging the gap between GNN research and real-world deployment requires co-design across algorithms, systems, and hardware. Future work should focus on developing RAN-compliant integration pipelines, efficient graph construction mechanisms, lifecycle-aware learning frameworks, and hardware-optimized GNN models that can operate reliably under real-time constraints.

### E. Mobility-Aware and Topology-Adaptive GNNs

A central open problem for AI-native 6G networks is the systematic treatment of mobility and topology evolution within graph-based learning. Existing dynamic GNNs provide promising building blocks, yet there is limited understanding of how to jointly model fast physical-layer dynamics (e.g., link intermittency, beam realignment) and slower network-layer events (e.g., handovers, routing updates) within a unified graph representation. Future work should investigate meta-learning, continual learning, and uncertainty-aware GNNs that can rapidly adapt to new or rarely seen topologies with minimal data, while explicitly accounting for the stochastic nature of mobility. Establishing common benchmarks and simulation environments that combine realistic mobility models (vehicular, UAV, LEO constellations) with dynamic graph datasets would be essential to assess and compare the robustness of topology-adaptive GNN solutions for 6G.

### F. Other Gaps–Benchmarking and Standardization

A persistent gap in the field is the absence of standardised benchmarks and evaluation frameworks for AI–GNN solutions in wireless networks. Unlike computer vision or Natural Language Processing (NLP), where mature datasets enable systematic comparison, the wireless community lacks unified scenarios for assessing AI–GNN performance. Recently, several wireless-specific resources have emerged, such as DeepMIMO, Raymobtime, and DeepSense-6G for channel, mobility, and sensing-rich data; and open RAN platforms including OpenRAN Gym, Colosseum datasets, and the O-RAN Software Community data releases for evaluating

AI-driven RAN control. In parallel, the GNNet challenge provides graph-structured topologies for benchmarking GNN inference on network graphs. These initiatives represent early steps toward reproducible and comparable evaluation of Hybrid GNN approaches. However, a comprehensive and widely adopted benchmarking suite, covering topology dynamics, multi-cell coordination, and real-time constraints, remains an open need. Likewise, industry bodies such as 3GPP and ITU are still defining how AI components, including GNN-based intelligence, should interface with network functions. Establishing common datasets, metrics, and standardized integration points will be essential to ensure interoperability and to enable fair, transparent validation of AI–GNN solutions as 6G matures.

## IX. Conclusion

AI-driven GNNs represent a convergence of two powerful paradigms uniquely suited for future wireless networks. Together, they offer 6G a way to learn and reason over the complex graphs that define network behaviour, from physical connectivity to network-awareness in communications. The journey to fully integrate AI–GNNs into 6G is just beginning, with formidable challenges in scalability, interpretability, and design still to be met. Yet, research momentum in the past two years (2024–2025) suggests these obstacles are surmountable. Key innovations around distributed learning, efficient architectures, and knowledge-infused models are paving the way for practical deployments. As these hurdles are overcome, the potential of GNN-Centric AI Architectures is set to reshape wireless architecture. We can envision 6G networks that operate as intelligent organisms sensing their environment, predicting changes, and coordinating responses in a distributed but unified manner. In such a paradigm, new applications flourish: communications become goal-oriented and efficient, networks heal and optimize themselves in real time, and the infrastructure transforms into a platform for intelligent services (from autonomous driving to immersive telepresence) far beyond the capabilities of today's networks. The road ahead will require deep collaboration between AI researchers and wireless engineers, but the destination, a 6G fabric that is truly intelligent by design, promises to fundamentally elevate the role of wireless connectivity in society. The interplay of GNN-Centric AI Architectures is thus not just an academic concept but a cornerstone of the 6G vision, one that will drive both technological and architectural innovation in the years to come.

## References


[1] F. Tariq, M. R. A. Khandaker, K.-K. Wong, M. A. Imran, M. Bennis, and M. Debbah, "A speculative study on 6G," *IEEE Wireless Commun.*, vol. 27, no. 4, pp. 118–125, Aug. 2020.

[2] M. H. Alsamh, A. Hawbani, S. Kumar, and S. H. Alsamhi, "Multisensory metaverse-6G: A new paradigm of commerce and education," *IEEE Access*, vol. 12, pp. 75657–75677, 2024.

[3] K. B. Letaief, W. Chen, Y. Shi, J. Zhang, and Y.-J.-A. Zhang, "The roadmap to 6G: AI empowered wireless networks," *IEEE Commun. Mag.*, vol. 57, no. 8, pp. 84–90, Aug. 2019.

[4] N. X. Tung et al., "Graph neural networks for next-generation-IoT: Recent advances and open challenges," *IEEE Commun. Surveys Tuts.*, vol. 28, pp. 2226–2262, 2026.

[5] M. Lee, G. Yu, H. Dai, and G. Y. Li, "Graph neural networks meet wireless communications: Motivation, applications, and future directions," *IEEE Wireless Commun.*, vol. 29, no. 5, pp. 12–19, Oct. 2022.

[6] P. Tam, S. Ros, I. Song, S. Kang, and S. Kim, "A survey of intelligent end-to-end networking solutions: Integrating graph neural networks and deep reinforcement learning approaches," *Electronics*, vol. 13, no. 5, p. 994, Mar. 2024.

[7] R. Liu, P. Xing, Z. Deng, A. Li, C. Guan, and H. Yu, "Federated graph neural networks: Overview, techniques, and challenges," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 36, no. 3, pp. 4279–4295, Mar. 2025.

[8] M. Nandan, S. Mitra, and D. De, "GraphXAI: A survey of graph neural networks (GNNs) for explainable AI (XAI)," *Neural Comput. Appl.*, vol. 37, no. 17, pp. 10949–11000, Jun. 2025.

[9] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Netw.*, vol. 34, no. 3, pp. 134–142, May 2020.

[10] A. Al-Ansi, A. M. Al-Ansi, A. Muthanna, I. A. Elgendy, and A. Koucheryavy, "Survey on intelligence edge computing in 6G: Characteristics, challenges, potential use cases, and market drivers," *Future Internet*, vol. 13, no. 5, p. 118, Apr. 2021.

[11] G. Fontanesi et al., "Artificial intelligence for satellite communication: A survey," *IEEE Commun. Surveys Tuts.*, vol. 28, pp. 1381–1435, 2025.

[12] K. Ntontin et al., "A vision, survey, and roadmap toward space communications in the 6G and beyond era," *Proc. IEEE*, vol. 113, no. 9, pp. 987–1023, Sep. 2025.

[13] R. Arshad and R. Muzzammel, "Realizing intelligence in 6G communications: A review," in *Proc. 2nd Int. Conf. Emerg. Trends Electr., Control, Telecommun. Eng. (ETECTE)*, Nov. 2023, pp. 1–5.

[14] H. M. F. Noman et al., "Machine learning empowered emerging wireless networks in 6G: Recent advancements, challenges and future trends," *IEEE Access*, vol. 11, pp. 83017–83051, 2023.

[15] H. Sun et al., "Advancing 6G: Survey for explainable AI on communications and network slicing," *IEEE Open J. Commun. Soc.*, vol. 6, pp. 1372–1412, 2025.

[16] Z. Chen, Z. Zhang, and Z. Yang, "Big AI models for 6G wireless networks: Opportunities, challenges, and research directions," *IEEE Wireless Commun.*, vol. 31, no. 5, pp. 164–172, Oct. 2024.

[17] L. Jiao et al., "Advanced deep learning models for 6G: Overview, opportunities, and challenges," *IEEE Access*, vol. 12, pp. 133245–133314, 2024.

[18] B. Zhao, J. Wu, Y. Ma, and C. Yang, "Meta-learning for wireless communications: A survey and a comparison to GNNs," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 1987–2015, 2024.

[19] Y. Song, T. Wang, P. Cai, S. K. Mondal, and J. P. Sahoo, "A comprehensive survey of few-shot learning: Evolution, applications, challenges, and opportunities," *ACM Comput. Surv.*, vol. 55, no. 13s, pp. 1–40, Dec. 2023.

[20] Q. Mao, F. Hu, and Q. Hao, "Deep learning for intelligent wireless networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 2595–2621, 4th Quart., 2018.

[21] S. Paul, "A comprehensive review on machine learning-based approaches for next generation wireless network," *Social Netw. Comput. Sci.*, vol. 5, no. 5, p. 468, Apr. 2024.

[22] Y.-F. Liu et al., "A survey of recent advances in optimization methods for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 11, pp. 2992–3031, Nov. 2024.

[23] C. Ssengonzi, O. P. Kogeda, and T. O. Olwal, "A survey of deep reinforcement learning application in 5G and beyond network slicing and virtualization," *Array*, vol. 14, Jul. 2022, Art. no. 100142.

[24] T. Li et al., "Applications of multi-agent reinforcement learning in future internet: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 2, pp. 1240–1279, 2nd Quart., 2022.

[25] S. K. Das, R. Mudi, M. S. Rahman, K. M. Rabie, and X. Li, "Federated reinforcement learning for wireless networks: Fundamentals, challenges and future research trends," *IEEE Open J. Veh. Technol.*, vol. 5, pp. 1400–1440, 2024.

[26] M. A. Ferrag et al., "Edge learning for 6G-enabled Internet of Things: A comprehensive survey of vulnerabilities, datasets, and defenses," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 4, pp. 2654–2713, 4th Quart., 2023.

[27] H. Pennanen, T. Hänninen, O. Tervo, A. Tölli, and M. Latva-Aho, "6G: The intelligent network of everything," *IEEE Access*, vol. 13, pp. 1319–1421, 2025.

[28] F. Tang, B. Mao, Y. Kawamoto, and N. Kato, "Survey on machine learning for intelligent end-to-end communication toward 6G: From network access, routing to traffic control and streaming adaption," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 3, pp. 1578–1598, 3rd Quart., 2021.

[29] P. Cardieri, "Modeling interference in wireless ad hoc networks," *IEEE Commun. Surveys Tuts.*, vol. 12, no. 4, pp. 551–572, 2010.

[30] P. H. Pathak and R. Dutta, "A survey of network design problems and joint design approaches in wireless mesh networks," *IEEE Commun. Surveys Tuts.*, vol. 13, no. 3, pp. 396–428, 3rd Quart., 2011.

[31] A. Majeed and I. Rauf, "Graph theory: A comprehensive survey about graph theory applications in computer science and social networks," *Inventions*, vol. 5, no. 1, p. 10, Feb. 2020.

[32] S. He et al., "An overview on the application of graph neural networks in wireless networks," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 2547–2565, 2021.

[33] W. Jiang, "Graph-based deep learning for communication networks: A survey," *Comput. Commun.*, vol. 185, pp. 40–54, Mar. 2022.

[34] P. Tam, I. Song, S. Kang, S. Ros, and S. Kim, "Graph neural networks for intelligent modelling in network management and orchestration: A survey on communications," *Electronics*, vol. 11, no. 20, p. 3371, Oct. 2022.

[35] Y. Li, S. Xie, Z. Wan, H. Lv, H. Song, and Z. Lv, "Graph-powered learning methods in the Internet of Things: A survey," *Mach. Learn. With Appl.*, vol. 11, Mar. 2023, Art. no. 100441.

[36] G. Dong et al., "Graph neural networks in IoT: A survey," *ACM Trans. Sensor Netw.*, vol. 19, no. 2, pp. 1–50, May 2023.

[37] A. Ivanov, K. Tonchev, V. Poulkov, A. Manolova, and N. N. Neshov, "Graph-based resource allocation for integrated space and terrestrial communications," *Sensors*, vol. 22, no. 15, p. 5778, Aug. 2022.

[38] J. Suárez-Varela et al., "Graph neural networks for communication networks: Context, use cases and opportunities," *IEEE Netw.*, vol. 37, no. 3, pp. 146–153, May 2023.

[39] Y. Dai et al., "A survey of graph-based resource management in wireless networks—Part I: Optimization approaches," *IEEE Trans. Cognit. Commun. Netw.*, vol. 11, no. 4, pp. 2078–2100, Aug. 2025.

[40] Y. Dai et al., "A survey of graph-based resource management in wireless networks—Part II: Learning approaches," *IEEE Trans. Cognit. Commun. Netw.*, vol. 11, no. 4, pp. 2101–2122, Aug. 2025.

[41] L. Zeng et al., "Edge graph intelligence: Reciprocally empowering edge networks with graph intelligence," *IEEE Commun. Surveys Tuts.*, vol. 27, no. 6, pp. 3417–3454, Dec. 2025.

[42] Y. Lu et al., "Agentic graph neural networks for wireless communications and networking toward edge general intelligence: A survey," *IEEE Commun. Surveys Tuts.*, vol. 28, pp. 4519–4554, 2026.

[43] S. Krishna Moorthy and J. Jagannath, "Survey of graph neural network for Internet of Things and NextG networks," *IEEE Commun. Surveys Tuts.*, vol. 28, pp. 1803–1844, 2026.

[44] D.-T. Ngo, O. Aouedi, K. Piamrat, T. Hassan, and P. Raipin-Parvédy, "Empowering digital twin for future networks with graph neural networks: Overview, enabling technologies, challenges, and opportunities," *Future Internet*, vol. 15, no. 12, p. 377, Nov. 2023.

[45] W. Chen et al., "5G-advanced toward 6G: Past, present, and future," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 6, pp. 1592–1619, Jun. 2023.

[46] F. Salahdine, T. Han, and N. Zhang, "5G, 6G, and beyond: Recent advances and future challenges," *Ann. Telecommun.*, vol. 78, nos. 9–10, pp. 525–549, Oct. 2023.

[47] A. Dogra, R. K. Jha, and S. Jain, "A survey on beyond 5G network with the advent of 6G: Architecture and emerging technologies," *IEEE Access*, vol. 9, pp. 67512–67547, 2021.

[48] *Framework and Overall Objectives of the Future Development of IMT for 2030 and Beyond*, document ITU-R, Recommendation ITU-R M.2160-0, Nov. 2023. [Online]. Available: https://www.itu.int/rec/R-REC-M.2160-0-202311-I/en

[49] C.-X. Wang et al., "On the road to 6G: Visions, requirements, key technologies, and testbeds," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 2, pp. 905–974, 2nd Quart. 2023.

[50] Q. Cui et al., "Overview of AI and communication for 6G network: Fundamentals, challenges, and future research opportunities," *Sci. China Inf. Sci.*, vol. 68, no. 7, Jul. 2025, Art. no. 171301.

[51] Y. Shen, J. Zhang, S. H. Song, and K. B. Letaief, "Graph neural networks for wireless communications: From theory to practice," *IEEE Trans. Wireless Commun.*, vol. 22, no. 5, pp. 3554–3569, May 2023.

[52] Z. Liu, J. Zhang, Y. Zhu, E. Shi, and B. Ai, "Robust multidimensional graph neural networks for signal processing in wireless communications with edge-graph information bottleneck," *IEEE Trans. Signal Process.*, vol. 73, pp. 2688–2703, 2025.

[53] J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl, "Neural message passing for quantum chemistry," in *Proc. 34th Int. Conf. Mach. Learning*, vol. 70, 2017, pp. 1263–1272.

[54] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2017.

[55] D. E. Heckerman and E. H. Shortliffe, "From certainty factors to belief networks," *Artif. Intell. Med.*, vol. 4, no. 1, pp. 35–52, Feb. 1992.

[56] H. Loeliger, "An introduction to factor graphs," *IEEE Signal Process. Mag.*, vol. 21, no. 1, pp. 28–41, Jan. 2004.

[57] B. Bollobás, *Modern Graph Theory* (Graduate Texts in Mathematics), vol. 184. New York, NY, USA: Springer, 1998.

[58] M. Gori, G. Monfardini, and F. Scarselli, "A new model for learning in graph domains," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, Mar. 2005, pp. 729–734.

[59] F. Scarselli, S. L. Yong, M. Gori, M. Hagenbuchner, A. C. Tsoi, and M. Maggini, "Graph neural networks for ranking web pages," in *Proc. IEEE/WIC/ACM Int. Conf. Web Intell. (WI)*, Sep. 2005, pp. 666–672.

[60] A. Pucci, M. Gori, M. Hagenbuchner, F. Scarselli, and A. C. Tsoi, "Investigation into the application of graph neural networks to large-scale recommender systems," *Syst. Sci.*, vol. 32, no. 4, pp. 17–26, 2006.

[61] P. Veličković, "Graph attention networks," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2018.

[62] C.-G. Li, J. Guo, and H.-G. Zhang, "Pruning neighborhood graph for geodesic distance based semi-supervised classification," in *Proc. Int. Conf. Comput. Intell. Secur. (CIS)*, Dec. 2007, pp. 428–432.

[63] Z. Zhang, P. Cui, J. Pei, X. Wang, and W. Zhu, "Eigen-GNN: A graph structure preserving plug-in for GNNs," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 3, pp. 2544–2555, Mar. 2023.

[64] Z. Xiang, T. Wang, and D. Wang, "Preserving node-level privacy in graph neural networks," in *Proc. IEEE Symp. Secur. Privacy (SP)*, May 2024, pp. 4714–4732.

[65] R. Dugyala et al., "Secure cloud computing: Leveraging GNN and leader K-means for intrusion detection optimization," *Sci. Rep.*, vol. 14, no. 1, p. 30906, Dec. 2024.

[66] M. Schlichtkrull, T. N. Kipf, P. Bloem, R. van den Berg, I. Titov, and M. Welling, "Modeling relational data with graph convolutional networks," in *Proc. Eur. Semantic Web Conf.*, 2018, pp. 593–607.

[67] Z. Sun et al., "Knowledge graph alignment network with gated multi-hop neighborhood aggregation," in *Proc. 34th AAAI Conf. Artif. Intell.*, 2020, pp. 222–229.

[68] Z.-Y. Li, Y.-Y. Zhou, and E.-H. He, "GraphSense: A self-aware dynamic graph learning networks for graph data over internet," *Appl. Intell.*, vol. 55, no. 1, p. 41, Jan. 2025.

[69] A. Pareja et al., "EvolveGCN: Evolving graph convolutional networks for dynamic graphs," in *Proc. AAAI Conf. Artif. Intell.*, Apr. 2020, vol. 34, no. 4, pp. 5363–5370.

[70] Y. Mo, X. Wang, S. Fan, and C. Shi, "Graph contrastive invariant learning from the causal perspective," in *Proc. AAAI Conf. Artif. Intell.*, 2024, vol. 38, no. 8, pp. 8904–8912.

[71] T. Lin et al., "Unraveling and mitigating endogenous task-oriented spurious correlations in ego-graphs via automated counterfactual contrastive learning," *Expert Syst. Appl.*, vol. 266, Mar. 2025, Art. no. 126015.

[72] D. Xu, C. Ruan, E. Körpeoğlu, S. Kumar, and K. Achan, "Inductive representation learning on temporal graphs," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2020.

[73] A. Sankar, Y. Wu, Z. Gouda, Y. Qi, and L. Jure, "DySAT: Dynamic self-attention network for temporal graphs," in *Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2020, pp. 6–10.

[74] L. Zhao et al., "T-GCN: A temporal graph convolutional network for traffic prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 9, pp. 3848–3858, Sep. 2020.

[75] X. Hao, C. She, P. Lep Yeoh, Y. Liu, B. Vucetic, and Y. Li, "Hybrid-task meta-learning: A GNN approach for scalable and transferable bandwidth allocation," *IEEE Trans. Wireless Commun.*, vol. 23, no. 12, pp. 19820–19835, Dec. 2024.

[76] R. Ying, D. Bourgeois, J. You, M. Žitnik, and J. Leskovec, "GNNExplainer: Generating explanations for graph neural networks," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2019.

[77] M. N. Vu and M. T. Thai, "PGM-explainer: Probabilistic graphical model explanations for graph neural networks," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 33, 2020, pp. 12225–12235.

[78] M. Eisen and A. Ribeiro, "Optimal wireless resource allocation with random edge graph neural networks," *IEEE Trans. Signal Process.*, vol. 68, pp. 2977–2991, 2020.

[79] T. Chen, X. Zhang, M. You, G. Zheng, and S. Lambotharan, "A GNN-based supervised learning framework for resource allocation in wireless IoT networks," *IEEE Internet Things J.*, vol. 9, no. 3, pp. 1712–1724, Feb. 2022.

[80] H. Yang et al., "Knowledge-driven resource allocation for wireless networks: A WMMSE unrolled graph neural network approach," *IEEE Internet Things J.*, vol. 11, no. 10, pp. 18902–18916, May 2024.

[81] D. Abode, R. Adeogun, and G. Berardinelli, "Power control for 6G in-factory subnetworks with partial channel information using graph neural networks," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 3120–3135, 2024.

[82] Y. Peng, J. Guo, and C. Yang, "Learning resource allocation policy: Vertex-GNN or edge-GNN?," *IEEE Trans. Mach. Learn. Commun. Netw.*, vol. 2, pp. 190–209, 2024.

[83] C. Zhang, J. Yin, and S. Deng, "Ensuring fairness in edge networks: A GNN-based media workload migration scheme with fairness guarantee," *IEEE Trans. Services Comput.*, vol. 17, no. 3, pp. 934–948, May 2024.

[84] Z. Zhao, G. Verma, C. Rao, A. Swami, and S. Segarra, "Link scheduling using graph neural networks," *IEEE Trans. Wireless Commun.*, vol. 22, no. 6, pp. 3997–4012, Jun. 2023.

[85] Z. Zhao, G. Verma, C. Rao, A. Swami, and S. Segarra, "Distributed scheduling using graph neural networks," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Jun. 2021, pp. 4720–4724.

[86] Z. Tian, X. Zhou, Z. Liao, M. Sun, and F. He, "GTSNet: A generalized traffic scheduler for time-sensitive networking based on graph neural network," *IEEE Trans. Ind. Informat.*, vol. 21, no. 1, pp. 208–217, Jan. 2025.

[87] Z. Wang, J. Guo, K. Parsons, Y. Nagai, T. Sumi, and P. Orlik, "Learning based routing link scheduling in heterogeneous wireless IoT networks," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Jun. 2024, pp. 1666–1672.

[88] J. Guo and C. Yang, "A size-generalizable graph neural network for learning multi-user multi-stream MIMO precoding," in *Proc. IEEE 34th Int. Workshop Mach. Learn. Signal Process. (MLSP)*, Sep. 2024, pp. 1–6.

[89] R. Wang, C. Yang, S. Han, J. Wu, S. Han, and X. Wang, "Learning end-to-end hybrid precoding for multi-user mmWave mobile system with GNNs," *IEEE Trans. Mach. Learn. Commun. Netw.*, vol. 2, pp. 978–993, 2024.

[90] C. He, Y. Lu, B. Ai, O. A. Dobre, Z. Ding, and D. Niyato, "ICGNN: Graph neural network enabled scalable beamforming for MISO interference channels," *IEEE Trans. Mobile Comput.*, vol. 24, no. 10, pp. 10778–10791, Oct. 2025.

[91] C. He, Y. Lu, R. Zhang, Y. Li, B. Ai, and D. Niyato, "GNN-enabled coordinated beamforming design for high speed railway communication systems," *IEEE Trans. Mobile Comput.*, pp. 1–13, Dec. 2025.

[92] Y. Li, Y. Lu, B. Ai, O. A. Dobre, Z. Ding, and D. Niyato, "GNN-based beamforming for sum-rate maximization in MU-MISO networks," *IEEE Trans. Wireless Commun.*, vol. 23, no. 8, pp. 9251–9264, Aug. 2024.

[93] Y. Li, Y. Lu, B. Ai, Z. Zhong, D. Niyato, and Z. Ding, "GNN-enabled max-min fair beamforming," *IEEE Trans. Veh. Technol.*, vol. 73, no. 8, pp. 12184–12188, Aug. 2024.

[94] Y. Hou et al., "NOMANet: A graph neural network enabled power allocation scheme for NOMA," *IEEE Trans. Veh. Technol.*, pp. 1–6, Jan. 2026.

[95] J. Guo and C. Yang, "Learning power allocation for multi-cell-multi-user systems with heterogeneous graph neural networks," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 884–897, Feb. 2022.

[96] B. Zhao, J. Guo, and C. Yang, "Understanding the performance of learning precoding policies with graph and convolutional neural networks," *IEEE Trans. Commun.*, vol. 72, no. 9, pp. 5657–5673, Sep. 2024.

[97] A. Chowdhury, G. Verma, A. Swami, and S. Segarra, "Deep graph unfolding for beamforming in MU-MIMO interference networks," *IEEE Trans. Wireless Commun.*, vol. 23, no. 5, pp. 4889–4903, May 2024.

[98] B. Xu, J. Zhang, Q. Lin, H. Xiao, Y.-C. Wu, and B. Ai, "Deep unfolding beamforming and power control designs for multi-port matching networks," *IEEE Trans. Wireless Commun.*, vol. 24, no. 2, pp. 1401–1414, Feb. 2025.

[99] S. Liu and C. Yang, "Learning user scheduling and hybrid precoding with sequential graph neural network," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2024, pp. 1–6.

[100] Y. Li, Y. Lu, G. Zhang, B. Ai, D. Niyato, and S. Cui, "Homogeneous and heterogeneous graph learning for hybrid beamforming in mmWave systems," *IEEE Trans. Wireless Commun.*, vol. 24, no. 10, pp. 8086–8100, Oct. 2025.

[101] Q. Wang, Y. Lu, W. Chen, B. Ai, Z. Zhong, and D. Niyato, "GNN-enabled optimization of placement and transmission design for UAV communications," *IEEE Trans. Veh. Technol.*, vol. 74, no. 4, pp. 6656–6661, Apr. 2025.

[102] A. Kosasih et al., "Graph neural network aided expectation propagation detector for MU-MIMO systems," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2022, pp. 1212–1217.

[103] H. He, X. Yu, J. Zhang, S. Song, and K. B. Letaief, "Message passing meets graph neural networks: A new paradigm for massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 23, no. 5, pp. 4709–4723, May 2024.

[104] W. Zhuang et al., "Approximate message passing-enhanced graph neural network for OTFS data detection," *IEEE Wireless Commun. Lett.*, vol. 13, no. 7, pp. 1913–1917, Jul. 2024.

[105] V. G. Satorras and M. Welling, "Neural enhanced belief propagation on factor graphs," in *Proc. Int. Conf. Artif. Intell. Statist. (AISTATS)*, 2020, pp. 685–693.

[106] Y. Liao, S. A. Hashemi, H. Yang, and J. M. Cioffi, "Scalable polar code construction for successive cancellation list decoding: A graph neural network-based approach," *IEEE Trans. Commun.*, vol. 71, no. 11, pp. 6231–6245, Nov. 2023.

[107] S. Cammerer, J. Hoydis, F. A. Aoudia, and A. Keller, "Graph neural networks for channel decoding," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2022, pp. 486–491.

[108] Z. Chen, L. Peng, and P.-H. Ho, "A novel GNN-based decoding scheme for sparse code multiple access (SCMA)," in *Proc. Int. Conf. Netw. Netw. Appl. (NaNA)*, Aug. 2024, pp. 160–165.

[109] S. Varadarajulu, V. M. Baeza, J. Querol, M. O. K. Mendonca, and S. Chatzinotas, "Graph neural network pooling for BCH channel decoding in software defined satellites," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Jun. 2024, pp. 1328–1333.

[110] Y. He, J. Gao, M. Fan, D. Han, Y. Lu, and Y. Qiao, "A&D graph-based graph neural network intrusion detection for in-vehicle controller area network," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, China, Aug. 2024, pp. 1281–1286.

[111] G. A. Al-Absi, Y. Fang, and A. A. Qaseem, "STC-GraphFormer: Graph spatial–temporal correlation transformer for in-vehicle network intrusion detection system," *Veh. Commun.*, vol. 51, Feb. 2025, Art. no. 100865.

[112] X. Han et al., "DE-GNN: Dual embedding with graph neural network for fine-grained encrypted traffic classification," *Comput. Netw.*, vol. 245, May 2024, Art. no. 110372.

[113] J. Yang, X. Jiang, Y. Lei, W. Liang, Z. Ma, and S. Li, "MTSecurity: Privacy-preserving malicious traffic classification using graph neural network and transformer," *IEEE Trans. Netw. Service Manage.*, vol. 21, no. 3, pp. 3583–3597, Jun. 2024.

[114] E. Akpaku, J. Chen, M. Ahmed, F. K. Agbenyegah, and W. L. Brown-Acquaye, "RAGN: Detecting unknown malicious network traffic using a robust adaptive graph neural network," *Comput. Netw.*, vol. 262, May 2025, Art. no. 111184.

[115] J. Lin et al., "Multi-scenario cellular KPI prediction based on spatiotemporal graph neural network," *IEEE Trans. Autom. Sci. Eng.*, vol. 22, pp. 5131–5142, 2025.

[116] B. Bertalanič, J. Hribar, and C. Fortuna, "Visibility graph-based wireless anomaly detection for digital twin edge networks," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 3050–3065, 2024.

[117] S. Das, N. NaderiAlizadeh, and A. Ribeiro, "State-augmented information routing in communication systems with graph neural networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2024, pp. 9181–9185.

[118] Z. Guo, F. Li, J. Shen, T. Xie, S. Jiang, and X. Wang, "ConfigReco: Network configuration recommendation with graph neural networks," *IEEE Netw.*, vol. 38, no. 1, pp. 7–14, Jan. 2024.

[119] Y. Ping, K. Xie, X. Huang, C. Li, and Y. Zhang, "GNN-based QoE optimization for dependent task scheduling in edge-cloud computing network," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2024, pp. 1–6.

[120] M. Mohsenivatani, S. Ali, V. Ranasinghe, N. Rajatheva, and M. Latva-Aho, "Graph representation learning for wireless communications," *IEEE Commun. Mag.*, vol. 62, no. 1, pp. 141–147, Jan. 2024.

[121] N. Hello, P. Di Lorenzo, and E. C. Strinati, "Semantic communication enhanced by knowledge graph representation learning," in *Proc. IEEE 25th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Sep. 2024, pp. 876–880.

[122] X. S. Xie, Y. Lu, and Z. Ding, "Graph neural network enabled pinching antennas," *IEEE Wireless Commun. Lett.*, vol. 14, no. 9, pp. 2982–2986, Jul. 2025.

[123] Z. Wang, Y. Zhou, Y. Zou, Q. An, Y. Shi, and M. Bennis, "A graph neural network learning approach to optimize RIS-assisted federated learning," *IEEE Trans. Wireless Commun.*, vol. 22, no. 9, pp. 6092–6106, Sep. 2023.

[124] H. Tang, J. Zhang, Z. Zhao, H. Wu, H. Sun, and P. Jiao, "Joint optimization based on two-phase GNN in RIS- and DF-assisted MISO systems with fine-grained rate demands," *IEEE Trans. Wireless Commun.*, vol. 24, no. 12, pp. 9989–10002, Dec. 2025.

[125] X. Li, M. Chen, Y. Hu, Z. Zhang, D. Liu, and S. Mao, "Jointly optimizing terahertz based sensing and communications in vehicular networks: A dynamic graph neural network approach," *IEEE Trans. Wireless Commun.*, vol. 23, no. 10, pp. 12917–12932, Oct. 2024.

[126] J. Wu, Z. Wang, Y.-F. Liu, and F. Liu, "Efficient global algorithms for transmit beamforming design in ISAC systems," *IEEE Trans. Signal Process.*, vol. 72, pp. 4493–4508, 2024.

[127] Q. Xue, H. Wang, H. Lv, and D. He, "High-precision indoor positioning via 5G NR: An interpretable GNN-based method," in *Proc. IEEE Int. Conf. Metaverse Comput., Netw., Appl. (MetaCom)*, Aug. 2024, pp. 265–272.

[128] K. Han, S. M. Yu, S.-L. Kim, and S.-W. Ko, "Mobility-induced graph learning for WiFi positioning," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 9, pp. 2487–2502, Sep. 2024.

[129] S. He, B. Yang, T. Liu, and J. Li, "Graph network-based UWB localization via learning spatial–temporal and geometric features," *IEEE Commun. Lett.*, vol. 29, no. 4, pp. 784–788, Apr. 2025.

[130] X. Li et al., "RadioGAT: A joint model-based and data-driven framework for multi-band radiomap reconstruction via graph attention networks," *IEEE Trans. Wireless Commun.*, vol. 23, no. 11, pp. 17777–17792, Nov. 2024.

[131] A. Bufort, L. Lebocq, and S. Cathabard, "Data-driven radio propagation modeling using graph neural networks," 2025, *arXiv:2501.06236*.

[132] J.-W. Lee et al., "Handover strategy for LEO satellite communication using graph neural network," *ICT Exp.*, vol. 11, no. 2, pp. 239–244, Apr. 2025.

[133] Y. Huang et al., "A GNN-enabled multipath routing algorithm for spatial–temporal varying LEO satellite networks," *IEEE Trans. Veh. Technol.*, vol. 73, no. 4, pp. 5454–5468, Apr. 2024.

[134] D. Zeng, W. Liu, W. Chen, L. Zhou, M. Zhang, and H. Qu, "Substructure aware graph neural networks," in *Proc. AAAI Conf. Artif. Intell.*, 2023, vol. 37, no. 9, pp. 11129–11137.

[135] Y. Zhang, Y. Gong, L. Fan, Y. Wang, Z. Han, and Y. Guo, "Quantum-assisted joint virtual network function deployment and maximum flow routing for space information networks," *IEEE Trans. Mobile Comput.*, vol. 24, no. 2, pp. 830–844, Feb. 2025.

[136] N. Heydaribeni, X. Zhan, R. Zhang, T. Eliassi-Rad, and F. Koushanfar, "Distributed constrained combinatorial optimization leveraging hypergraph neural networks," *Nature Mach. Intell.*, vol. 6, no. 6, pp. 664–672, May 2024.

[137] J. Zhang et al., "Multi-agent reinforcement learning in wireless distributed networks for 6G," 2025, *arXiv:2502.05812*.

[138] S. Munikoti, D. Agarwal, L. Das, M. Halappanavar, and B. Natarajan, "Challenges and opportunities in deep reinforcement learning with graph neural networks: A comprehensive review of algorithms and applications," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 11, pp. 15051–15071, Nov. 2024.

[139] Z. Liu et al., "Graph neural network meets multi-agent reinforcement learning: Fundamentals, applications, and future directions," *IEEE Wireless Commun.*, vol. 31, no. 6, pp. 39–47, Dec. 2024.

[140] L. M. Amorosa, M. Skocaj, R. Verdone, and D. Gündüz, "Multi-agent reinforcement learning for power control in wireless networks via adaptive graphs," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2024, pp. 2968–2973.

[141] Z. Liu, J. Zhang, Y. Zhu, E. Shi, and B. Ai, "Mobile cell-free massive MIMO with multi-agent reinforcement learning: A scalable framework," *IEEE Trans. Wireless Commun.*, vol. 24, no. 3, pp. 1817–1831, Mar. 2025.

[142] Y. Du et al., "Distributed multi-UAV trajectory planning for downlink transmission: A GNN-enhanced DRL approach," *IEEE Wireless Commun. Lett.*, vol. 13, no. 12, pp. 3578–3582, Dec. 2024.

[143] Z. He, L. Wang, H. Ye, G. Y. Li, and B.-H.-F. Juang, "Resource allocation based on graph neural networks in vehicular communications," in *Proc. GLOBECOM IEEE Global Commun. Conf.*, Dec. 2020, pp. 1–5.

[144] Z. Wang, M. Bennis, and Y. Zhou, "Graph attention-based MADRL for access control and resource allocation in wireless networked control systems," *IEEE Trans. Wireless Commun.*, vol. 23, no. 11, pp. 16076–16090, Nov. 2024.

[145] X. Zhang et al., "Decentralized routing and radio resource allocation in wireless ad hoc networks via graph reinforcement learning," *IEEE Trans. Cognit. Commun. Netw.*, vol. 10, no. 3, pp. 1146–1159, Jun. 2024.

[146] J. Zhao and C. Yang, "Graph reinforcement learning with relational priors for predictive power allocation," *Sci. China Inf. Sci.*, vol. 68, no. 2, pp. 1–18, Feb. 2025.

[147] Z. Wang, T. Gong, and S. H. Huang, "Graph neural network with soft actor-critic and attention-based large model for intelligent edge routing in consumer Internet of Things," *IEEE Trans. Consum. Electron.*, vol. 71, no. 3, pp. 9061–9074, Aug. 2025.

[148] Z. Wang, B. Yi, S. Kumari, C. M. Chen, and M. J. F. Alenazi, "Graph convolutional networks and deep reinforcement learning for intelligent edge routing in IoT environment," *Comput. Commun.*, vol. 232, Feb. 2025, Art. no. 108050.

[149] W. G. Negassa, D. J. Gelmecha, R. S. Singh, and D. S. Rathee, "Hybrid machine learning-based 3-dimensional UAV node localization for UAV-assisted wireless networks," *Cognit. Robot.*, vol. 5, pp. 61–76, Jan. 2025.

[150] B. Hazarika et al., "Generative AI-augmented graph reinforcement learning for adaptive UAV swarm optimization," *IEEE Internet Things J.*, vol. 12, no. 8, pp. 9508–9524, Apr. 2025.

[151] L. Bai, H. Ma, Y. Jiang, Z. Yin, H. Wan, and H. Wang, "GRL-RR: A graph reinforcement learning-based resilient routing framework for software-defined LEO mega-constellations," *Comput. Netw.*, vol. 259, Mar. 2025, Art. no. 111089.

[152] S. Zhang et al., "GRLR: Routing with graph neural network and reinforcement learning for mega LEO satellite constellations," *IEEE Trans. Veh. Technol.*, vol. 74, no. 2, pp. 3225–3237, Feb. 2025.

[153] P. Xu, M. Feng, J. Zhou, L. Xiao, P. Xiao, and T. Jiang, "Inter-satellite routing for LEO satellite networks: A GNN and DRL integrated approach," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Aug. 2024, pp. 1346–1351.

[154] S. Sugapriya and R. Vijayabhasker, "Improved deep reinforcement learning with faster graph recurrent convolutional neural network-enabled adaptive network slicing for tailored service delivery in NextGen networks," *Int. J. Commun. Syst.*, vol. 38, no. 2, p. 6049, Jan. 2025.

[155] Y. Tan, J. Liu, and J. Wang, "5G end-to-end slice embedding based on heterogeneous graph neural network and reinforcement learning," *IEEE Trans. Cognit. Commun. Netw.*, vol. 10, no. 3, pp. 1119–1131, Jun. 2024.

[156] D. Wang, Y. Bai, and B. Song, "A knowledge graph-based reinforcement learning approach for cooperative caching in MEC-enabled heterogeneous networks," *Digit. Commun. Netw.*, vol. 11, no. 4, pp. 1237–1245, Aug. 2025.

[157] J. Shen, Y. Lin, Y. Zhang, W. Zhang, F. Shu, and J. Li, "Content caching-assisted vehicular edge computing using multi-agent graph attention reinforcement learning," *IEEE Trans. Veh. Technol.*, vol. 74, no. 2, pp. 3509–3514, Feb. 2025.

[158] S. Khan, M. Avgeris, J. Gascon-Samson, and A. Leivadeas, "EMDTORA: Energy-aware multi-user dependent task offloading and resource allocation in MEC using graph-enabled DRL," *IEEE Trans. Green Commun. Netw.*, vol. 9, no. 3, pp. 1453–1469, Sep. 2025.

[159] H. Wu, L. Tian, H. Tang, R. Li, and P. Jiao, "Graph convolutional reinforcement learning-guided joint trajectory optimization and task offloading for aerial edge computing," *IEEE Trans. Intell. Transp. Syst.*, vol. 26, no. 10, pp. 17487–17498, Oct. 2025.

[160] H. Yonghui, J. Zuodong, Q. Peng, and T. Dan, "Task offloading and resource allocation in NOMA-VEC: A multi-agent deep graph reinforcement learning algorithm," *China Commun.*, vol. 21, no. 8, pp. 79–88, Aug. 2024.

[161] Z. Cao, X. Deng, S. Yue, P. Jiang, J. Ren, and J. Gui, "Dependent task offloading in edge computing using GNN and deep reinforcement learning," *IEEE Internet Things J.*, vol. 11, no. 12, pp. 21632–21646, Jun. 2024.

[162] F. Yang et al., "Mutual-interference-aware throughput enhancement in massive IoT: A graph reinforcement learning framework," *IEEE Internet Things J.*, vol. 11, no. 18, pp. 30341–30353, Sep. 2024.

[163] F. Li, J. Yang, K.-Y. Lam, B. Shen, and G. Wei, "Dynamic spectrum access for Internet-of-Things with joint GNN and DQN," *Ad Hoc Netw.*, vol. 163, Oct. 2024, Art. no. 103596.

[164] W. Sun, Y. Zou, N. Guan, X. Zhang, G. Du, and Y. Wen, "Graph attention network–based deep reinforcement learning scheduling framework for in-vehicle time-sensitive networking," *IEEE Trans. Ind. Informat.*, vol. 20, no. 7, pp. 9825–9836, Jul. 2024.

[165] X. Zhao, H. Yang, and M. Li, "Graph reinforcement learning based multi-hotspot region UAV dynamic scheduling in mobile edge computing," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2024, pp. 1–6.

[166] Z. Shan, X. Yi, L. Liang, C.-S. Liao, and S. Jin, "GRLinQ: A distributed link scheduling mechanism with graph reinforcement learning," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2024, pp. 2436–2441.

[167] M. Ding, Y. Guo, Z. Huang, B. Lin, and H. Luo, "GROM: A generalized routing optimization method with graph neural network and deep reinforcement learning," *J. Netw. Comput. Appl.*, vol. 229, Sep. 2024, Art. no. 103927.

[168] J. Wu, Q. Guangqiu, C. Wu, W. Jiang, and J. Jiahe, "Federated learning for network attack detection using attention-based graph neural networks," *Sci. Rep.*, vol. 14, no. 1, p. 19088, 2024.

[169] M. Lee, G. Yu, and H. Dai, "Privacy-preserving decentralized inference with graph neural networks in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 23, no. 1, pp. 543–558, Jan. 2024.

[170] L. Zhong, Z. Chen, H. Cheng, and J. Li, "Lightweight federated graph learning for accelerating classification inference in UAV-assisted MEC systems," *IEEE Internet Things J.*, vol. 11, no. 12, pp. 21180–21190, Jun. 2024.

[171] G. Huang, X. Chen, Q. Wu, J. Li, and Q. Huang, "Joint client and cross-client edge selection for cost-efficient federated learning of graph convolutional networks," *IEEE Trans. Netw.*, vol. 33, no. 1, pp. 286–298, Feb. 2024.

[172] N. Yang, S. Wang, M. Chen, C. G. Brinton, and C. Yin, "Energy efficient collaborative federated learning design: A graph neural network based approach," in *Proc. GLOBECOM IEEE Global Commun. Conf.*, Dec. 2023, pp. 164–169.

[173] C. Meng, M. Tang, M. Setayesh, and V. W. S. Wong, "GNN-based neighbor selection and resource allocation for decentralized federated learning," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2023, pp. 1223–1228.

[174] H. Fereidooni et al., "SAFELearn: Secure aggregation for private FEderated learning," in *Proc. IEEE Secur. Privacy Workshops (SPW)*, May 2021, pp. 56–62.

[175] Z. Zhang, J. Li, S. Yu, and C. Makaya, "SAFELearning: Secure aggregation in federated learning with backdoor detectability," *IEEE Trans. Inf. Forensics Security*, vol. 18, pp. 3289–3304, 2023.

[176] K. Wei et al., "Federated learning with differential privacy: Algorithms and performance analysis," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 3454–3469, 2020.

[177] E. Rizk and A. H. Sayed, "A graph federated architecture with privacy preserving learning," in *Proc. IEEE 22nd Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Sep. 2021, pp. 131–135.

[178] G. Zhao, Y. Huang, and C. H. Tsai, "FedGSL: Federated graph structure learning for local subgraph augmentation," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2022, pp. 818–824.

[179] Y. Zhang et al., "A survey on privacy in graph neural networks: Attacks, preservation, and applications," *IEEE Trans. Knowl. Data Eng.*, vol. 36, no. 12, pp. 7497–7515, Dec. 2024.

[180] E. Yu, S. Liu, Q. Li, H. Chen, H. V. Poor, and S. Shamai (Shitz), "Graph-based joint client clustering and resource allocation for wireless distributed learning: A new hierarchical federated learning framework with non-IID data," *IEEE Trans. Mobile Comput.*, vol. 24, no. 5, pp. 3579–3596, May 2025.

[181] Z. Wang, Y. Zhou, and Y. Shi, "Over-the-air federated graph learning," *IEEE Trans. Wireless Commun.*, vol. 23, no. 12, pp. 18669–18683, Dec. 2024.

[182] K. T. Pauu, J. Wu, Y. Fan, Q. Pan, and M.-V. Maka, "Differential privacy and blockchain-empowered decentralized graph federated learning-enabled UAVs for disaster response," *IEEE Internet Things J.*, vol. 11, no. 12, pp. 20930–20947, Nov. 2023.

[183] S. Krishnan, J. Park, S. Sagar, G. Sherman, B. Campbell, and J. Choi, "Federated graph learning for low probability of detection in wireless ad-hoc networks," in *Proc. IEEE Stat. Signal Process. Workshop (SSP)*, Jul. 2023, pp. 66–70.

[184] Y. Etiabi et al., "MetaGraphLoc: A graph-based meta-learning scheme for indoor localization via sensor fusion," 2024, *arXiv:2411.17781*.

[185] Y. Cai, P. Cheng, Z. Chen, W. Xiang, B. Vucetic, and Y. Li, "Graphic deep reinforcement learning for dynamic resource allocation in space-air-ground integrated networks," *IEEE J. Sel. Areas Commun.*, vol. 43, no. 1, pp. 334–349, Jan. 2025.

[186] Q. Hou, M. Lee, G. Yu, and Y. Cai, "Meta-gating framework for fast and continuous resource optimization in dynamic wireless environments," *IEEE Trans. Commun.*, vol. 71, no. 9, pp. 5259–5273, Sep. 2023.

[187] M. Chen et al., "Cross-modal graph semantic communication assisted by generative AI in the metaverse for 6G," *Research*, vol. 7, p. 342, Jan. 2024.

[188] H. Chen, C. Xu, L. Zheng, Q. Zhang, and X. Lin, "Diffusion-based graph generative methods," *IEEE Trans. Knowl. Data Eng.*, vol. 36, no. 12, pp. 7954–7972, Dec. 2024.

[189] D. Fan et al., "Generative diffusion models for wireless networks: Fundamental, architecture, and state-of-the-art," *IEEE Commun. Surveys Tuts.*, early access, Mar. 5, 2026, doi: 10.1109/COMST.2026.3671110.

[190] Y. B. Uslu, S. Hadou, S. S. Bidokhti, and A. Ribeiro, "GNN-parametrized diffusion policies for wireless resource allocation," in *Proc. NeurIPS Workshop New Perspect. Graph Mach. Learn.*, 2025.

[191] Z. Li, L. Xia, H. Hua, S. Zhang, S. Wang, and C. Huang, "DiffGraph: Heterogeneous graph diffusion model," in *Proc. 18th ACM Int. Conf. Web Search Data Mining*, Mar. 2025, pp. 40–49.

[192] R. Liang et al., "GDSG: Graph diffusion-based solution generator for optimization problems in MEC networks," *IEEE Trans. Mobile Comput.*, vol. 24, no. 10, pp. 10264–10277, Oct. 2025.

[193] R. A. Jacobs, M. I. Jordan, S. J. Nowlan, and G. E. Hinton, "Adaptive mixtures of local experts," *Neural Comput.*, vol. 3, no. 1, pp. 79–87, Mar. 1991.

[194] N. Shazeer et al., "Outrageously large neural networks: The sparsely-gated mixture-of-experts layer," 2017, *arXiv:1701.06538*.

[195] Y. Han et al., "Graph mixture of experts: Learning on large-scale graphs with explicit diversity modeling," in *Proc. Adv. Neural Inf. Process. Syst.*, 2023, pp. 50825–50837.

[196] Z. Chen et al., "Mixture-of-experts with expert choice routing," in *Proc. Adv. Neural Inf. Process. Syst.*, 2022, pp. 7103–7114.

[197] L. Wang, H. Gao, C. Zhao, X. Sun, and D. Dai, "Auxiliary-loss-free load balancing strategy for mixture-of-experts," 2024, *arXiv:2408.15664*.

[198] D. Zheng et al., "DistDGL: Distributed graph neural network training for billion-scale graphs," in *Proc. IEEE/ACM 10th Workshop Irregular Appl., Archit. Algorithms (IA)*, Nov. 2020, pp. 36–44.

[199] Z. Yao et al., "DA-MoE: Addressing depth-sensitivity in graph-level analysis through mixture of experts," *Neural Netw.*, vol. 194, Feb. 2026, Art. no. 108064.

[200] T. Nguyen, N. N. Tran, K. Nguyen, and R. G. Baraniuk, "Improving routing in sparse mixture of experts with graph of tokens," 2025, *arXiv:2505.00792*.

[201] H. Liu, Y. Zhang, X. Wang, B. Wang, and Y. Yu, "ST-MoE: Spatio-temporal mixture of experts for multivariate time series forecasting," in *Proc. 18th Int. Conf. Intell. Syst. Knowl. Eng. (ISKE)*, Nov. 2023, pp. 562–567.

[202] H. Kang, T. Kim, H. Yang, and H. Choo, "Movement attentive graph embedding for improving next point of access prediction," in *Proc. 18th Int. Conf. Ubiquitous Inf. Manage. Commun. (IMCOM)*, Jan. 2024, pp. 1–3.

[203] Y. Matsuo et al., "Deep learning, reinforcement learning, and world models," *Neural Netw.*, vol. 152, pp. 267–275, Aug. 2022.

[204] C. Kurisummoottil Thomas, W. Saad, and Y. Xiao, "Causal semantic communication for digital twins: A generalizable imitation learning approach," *IEEE J. Sel. Areas Inf. Theory*, vol. 4, pp. 698–717, 2023.

[205] A. Khanna, G. Anjali, N. K. Verma, and K. J. Naik, "A GRL-aided federated graph reinforcement learning approach for enhanced file caching in mobile edge computing," *Computing*, vol. 107, no. 1, p. 40, Jan. 2025.

[206] T. Wang et al., "FlagVNE: A flexible and generalizable reinforcement learning framework for network resource allocation," in *Proc. Int. Joint Conf. Artif. Intell. (IJCAI)*, 2024, pp. 2406–2414.

[207] A. Mesodiakaki et al., "Online energy-efficient resource allocation in integrated terrestrial and satellite 6G networks," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2024, pp. 2083–2088.

[208] J. Zhu, X. Lin, Y. Qi, G. Li, C. Zhang, and J. Li, "Local differential private spatio-temporal dynamic graph learning for wireless social networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2024, pp. 1–6.

[209] S. Deka, K. Deka, N. T. Nguyen, S. Sharma, V. Bhatia, and N. Rajatheva, "Comprehensive review of deep unfolding techniques for next-generation wireless communication systems," *IEEE Internet Things J.*, vol. 13, no. 6, pp. 10379–10406, Mar. 2026.

[210] C. Miao, J. Zhang, J. Huang, and J. Yang, "A graph neural network power allocation algorithm based on fully unrolled WMMSE," in *Proc. IEEE 25th Int. Symp. World Wireless, Mobile Multimedia Netw. (WoWMoM)*, Jun. 2024, pp. 138–146.

[211] B. Li, G. Verma, and S. Segarra, "Graph-based algorithm unfolding for energy-aware power allocation in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 22, no. 2, pp. 1359–1373, Feb. 2023.

[212] J. Liu and X. Wang, "Tanner-graph-based massive multiple access—Transmission and decoding schemes," *IEEE Trans. Wireless Commun.*, vol. 22, no. 11, pp. 8468–8482, Nov. 2023.

[213] N. Li, "Towards AI-native RAN: An operator's perspective of 6G day 1 standardization," 2025, *arXiv:2507.08403*.

[214] X. Lin, "3GPP evolution from 5G to 6G: A 10-year retrospective," *Telecom*, vol. 6, no. 2, p. 32, May 2025.

[215] X. Li, J. Zhao, and M. Peng, "A survey on network slicing: Enabling technologies and solutions," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 1, pp. 472–507, 4th Quart., 2021.

[216] A. Moayyedi, M. Ahmadi, M. A. Salahuddin, R. Boutaba, and A. Saleh, "Generalizable GNN-based 5G RAN/MEC slicing and admission control in metropolitan networks," in *Proc. NOMS IEEE/IFIP Netw. Operations Manage. Symp.*, May 2023, pp. 1–9.

[217] A. Zolghadr, J. F. Santos, L. A. DaSilva, and J. Kibilda, "Learning and reconstructing conflicts in O-RAN: A graph neural network approach," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2025, pp. 1–6.

[218] M. A. Shami, J. Yan, and E. T. Fapi, "O-RAN xApps conflict prediction using graph convolutional networks," *IEEE Trans. Mobile Comput.*, pp. 1–17, Feb. 2026.

[219] P. Sharma, S. Sun, S. Deshpande, A. Stavrou, and H. Wang, "Towards xApp conflict evaluation with explainable machine learning and causal inference in O-RAN," 2025, *arXiv:2510.13031*.

[220] S. Piroti, A. Chawla, and T. Zanouda, "Mobile network configuration recommendation using deep generative graph neural network," *IEEE Netw. Lett.*, vol. 6, no. 3, pp. 179–182, Sep. 2024.

[221] A. S. Abdalla and V. Marojevic, "The role of generative artificial intelligence in testing and automation of open RAN systems: GNG-OTA framework," in *Proc. IEEE Int. Conf. Omni-layer Intell. Syst. (COINS)*, Aug. 2025, pp. 1–5.

[222] A. Hasan, C. Boeira, K. Papry, Y. Ju, Z. Zhu, and I. Haque, "Root cause analysis of anomalies in 5G RAN using graph neural network and transformer," 2024, *arXiv:2406.15638*.

[223] E. Björnson and L. Sanguinetti, "Scalable cell-free massive MIMO systems," *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 4247–4261, Jul. 2020.

[224] U. Demirhan and A. Alkhateeb, "Learning beamforming in cell-free massive MIMO ISAC systems," in *Proc. IEEE 25th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Sep. 2024, pp. 326–330.

[225] Z. Wang and V. Wong, "Heterogeneous graph neural network for cooperative ISAC beamforming in cell-free MIMO systems," in *Proc. 30th Annu. Int. Conf. Mobile Comput. Netw.*, Dec. 2024, pp. 2161–2172.

[226] B. Parlier, L. Salaün, and H. Yang, "Learning optimal linear precoding for cell-free massive MIMO with GNN," in *Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discov. Databases (ECML PKDD)*, 2024, pp. 72–88.

[227] B. Li, L.-L. Yang, R. G. Maunder, S. Sun, and P. Xiao, "Heterogeneous graph neural network for power allocation in multicarrier-division duplex cell-free massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 23, no. 2, pp. 962–977, Feb. 2024.

[228] B. Peng et al., "Distributed combinatorial optimization of downlink user assignment in mmWave cell-free massive MIMO using graph neural networks," in *Proc. GLOBECOM IEEE Global Commun. Conf.*, Dec. 2024, pp. 462–468.

[229] X. Yan, Z. Wang, Y. Jia, Z. Zhang, and Y. Huang, "Access point selection and beamforming design for cell-free network: From fractional programming to GNN," *IEEE Trans. Wireless Commun.*, vol. 23, no. 8, pp. 9345–9360, Aug. 2024.

[230] T. Van Chien, H. A. Le, T. H. Tung, H. Q. Ngo, and S. Chatzinotas, "Joint power allocation and user scheduling in integrated satellite–terrestrial cell-free massive MIMO IoT systems," *IEEE Internet Things J.*, vol. 11, no. 20, pp. 32883–32900, Oct. 2024.

[231] A. Mehrabian and V. W. S. Wong, "Joint spectrum, precoding, and phase shifts design for RIS-aided multiuser MIMO THz systems," *IEEE Trans. Commun.*, vol. 72, no. 8, pp. 5087–5101, Aug. 2024.

[232] B.-J. Chen, R. Y. Chang, F.-T. Chien, and H. V. Poor, "Energy-efficient multi-RIS-aided rate-splitting multiple access: A graph neural network approach," *IEEE Wireless Commun. Lett.*, vol. 13, no. 7, pp. 2003–2007, Jul. 2024.

[233] H. A. Le, T. Van Chien, and W. Choi, "Graph neural network-based active and passive beamforming for distributed STAR-RIS-assisted multi-user MISO systems," *IEEE Trans. Commun.*, vol. 73, no. 10, pp. 9299–9312, Oct. 2025.

[234] J. Yang, J. Xu, Y. Zhang, H. Zheng, and T. Zhang, "Graph attention network-based precoding for reconfigurable intelligent surfaces aided wireless communication systems," *IEEE Trans. Veh. Technol.*, vol. 73, no. 6, pp. 9098–9102, Jun. 2024.

[235] M. Dai, T. Gong, and S. Zhang, "Graph convolutional network-based channel extrapolation for hybrid RIS-aided communication," *IEEE Wireless Commun. Lett.*, vol. 13, no. 12, pp. 3558–3562, Dec. 2024.

[236] S. Lyu, L. Peng, and S. Y. Chang, "Investigating large-scale RIS-assisted wireless communications using GNN," *IEEE Trans. Consum. Electron.*, vol. 70, no. 1, pp. 811–818, Feb. 2024.

[237] S. S. Gill et al., "Modern computing: Vision and challenges," *Telemat. Inform. Rep.*, vol. 13, Mar. 2024, Art. no. 100116.

[238] V. Veeramachaneni, "Edge computing: Architecture, applications, and future challenges in a decentralized era," *Recent Trends Comput. Graph. Multimedia Technol.*, vol. 7, no. 1, pp. 8–23, 2025.

[239] V. P. Rekkas, S. Sotiroudis, P. Sarigiannidis, S. Wan, G. K. Karagiannidis, and S. K. Goudos, "Machine learning in beyond 5G/6G networks—State-of-the-art and future trends," *Electronics*, vol. 10, no. 22, p. 2786, Nov. 2021.

[240] Y. Wang, H. Wu, and R. Li, "Deep graph reinforcement learning for mobile edge computing: Challenges and solutions," *IEEE Netw.*, vol. 38, no. 5, pp. 314–323, Sep. 2024.

[241] T. Wang, X. Ouyang, D. Sun, Y. Chen, and H. Li, "Offloading strategy based on graph neural reinforcement learning in mobile edge computing," *Electronics*, vol. 13, no. 12, p. 2387, Jun. 2024.

[242] Y. Yang, L. Feng, Y. Sun, Y. Li, W. Li, and M. A. Imran, "Multi-cluster cooperative offloading for VR task: A MARL approach with graph embedding," *IEEE Trans. Mobile Comput.*, vol. 23, no. 9, pp. 8773–8788, Sep. 2024.

[243] A. Xu et al., "TransEdge: Task offloading with GNN and DRL in edge-computing-enabled transportation systems," *IEEE Internet Things J.*, vol. 11, no. 23, pp. 38151–38166, Dec. 2024.

[244] Q. Jiang et al., "Potential game based distributed IoV service offloading with graph attention networks in mobile edge computing," *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 9, pp. 10912–10925, Sep. 2024.

[245] C. Song, J. Wu, K. Xian, J. Huang, and L. Lu, "Spatio-temporal graph learning: Traffic flow prediction of mobile edge computing in 5G/6G vehicular networks," *Comput. Netw.*, vol. 252, Oct. 2024, Art. no. 110676.

[246] A. Ali et al., "A resource-aware multi-graph neural network for urban traffic flow prediction in multi-access edge computing systems," *IEEE Trans. Consum. Electron.*, vol. 70, no. 4, pp. 7252–7265, Nov. 2024.

[247] A. Asheralieva, D. Niyato, and Y. Miyanaga, "Efficient dynamic distributed resource slicing in 6G multi-access edge computing networks with online ADMM and message passing graph neural networks," *IEEE Trans. Mobile Comput.*, vol. 23, no. 4, pp. 2614–2638, Apr. 2024.

[248] F. Li, J. Lin, Y. Wang, Y. Du, and H. Han, "Distributed hierarchical temporal graph learning for communication-efficient high-dimensional industrial IoT modeling," *IEEE Internet Things J.*, vol. 11, no. 17, pp. 28578–28590, Sep. 2024.

[249] J. Graham, R. Dividino, and R. De Grande, "Dynamic network edge analysis for Internet of Vehicles with graph neural networks," in *Proc. 20th Int. Conf. Wireless Mobile Comput., Netw. Commun. (WiMob)*, Oct. 2024, pp. 371–376.

[250] M. Jia, J. Wu, Q. Guo, and Y. Yang, "Service-oriented SAGIN with pervasive intelligence for resource-constrained users," *IEEE Netw.*, vol. 38, no. 2, pp. 79–86, Mar. 2024.

[251] W. Wang, W. Ma, and K. Yan, "Trust-aware privacy-preserving QoS prediction with graph neural collaborative filtering for Internet of Things services," *Complex Intell. Syst.*, vol. 11, no. 4, pp. 1–18, Apr. 2025.

[252] M. Liu, H. Xu, Q. Z. Sheng, and Z. Wang, "QoSGNN: Boosting QoS prediction performance with graph neural networks," *IEEE Trans. Services Comput.*, vol. 17, no. 2, pp. 645–658, Mar. 2024.

[253] H. Wu, S. Tian, B. Jin, Y. Zhao, and L. Zhang, "Effective graph modeling and contrastive learning for time-aware QoS prediction," *IEEE Trans. Services Comput.*, vol. 17, no. 6, pp. 3513–3526, Nov. 2024.

[254] S. Kumar, S. Chattopadhyay, and C. Adak, "Anomaly resilient temporal QoS prediction using hypergraph convoluted transformer network," *IEEE Trans. Netw. Service Manag.*, early access, Mar. 2026.

[255] S. Hu et al., "GACL: Graph attention collaborative learning for temporal QoS prediction," *IEEE Trans. Netw. Service Manage.*, vol. 22, no. 4, pp. 3388–3402, Aug. 2025.

[256] Z. Niu, M. Roveri, and G. Casale, "ChainNet: A customized graph neural network model for loss-aware edge AI service deployment," in *Proc. 54th Annu. IEEE/IFIP Int. Conf. Dependable Syst. Netw. (DSN)*, Jun. 2024, pp. 238–251.

[257] K. Zacharopoulos, G. Koutroumpas, I. Arapakis, K. Georgopoulos, J. Khangosstar, and S. Ioannidis, "LightningNet: Distributed graph-based cellular network performance forecasting for the edge," 2024, *arXiv:2403.18810*.

[258] B. Liu, X. Chen, Q. Yuan, D. Li, and C. Gu, "TMC-GCN: Encrypted traffic mapping classification method based on graph convolutional networks," *Comput., Mater. Continua*, vol. 82, no. 2, pp. 3179–3201, 2025.

[259] R. Yu, X. Guo, P. Zhang, and K. Zhang, "HGNN-ETC: Higher-order graph neural network based on chronological relationships for encrypted traffic classification," *Comput., Mater. Continua*, vol. 81, no. 2, pp. 2643–2664, 2024.

[260] M. Liu, Q. Yang, W. Wang, and S. Liu, "TB-graph: Enhancing encrypted malicious traffic classification through relational graph attention networks," *Comput., Mater. Continua*, vol. 82, no. 2, pp. 2985–3004, 2025.

[261] Z. Li, H. Zhao, J. Zhao, Y. Jiang, and F. Bu, "SAT-Net: A staggered attention network using graph neural networks for encrypted traffic classification," *J. Netw. Comput. Appl.*, vol. 233, Jan. 2025, Art. no. 104069.

[262] H. Latif-Martínez, J. Suárez-Varela, A. Cabellos-Aparicio, and P. Barlet-Ros, "GAT-AD: Graph attention networks for contextual anomaly detection in network monitoring," *Comput. Ind. Eng.*, vol. 200, Feb. 2025, Art. no. 110830.

[263] B. K. Dhamala, B. R. Dawadi, P. Manzoni, and B. K. Acharya, "Performance evaluation of graph neural network-based RouteNet model with attention mechanism," *Future Internet*, vol. 16, no. 4, p. 116, Mar. 2024.

[264] H. Zhou, D. Yao, B. Chen, K. Yu, and X. Wu, "Graph convolution network based state space model for wireless traffic prediction," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2024, pp. 1–6.

[265] L. Li and K. Kim, "GTT-NTP: A graph convolutional networks-based network traffic prediction model," in *Proc. NOMS IEEE Netw. Oper. Manage. Symp.*, May 2024, pp. 1–7.

[266] T. Zhang, T. Luo, S. Huang, Y. Li, and X. Wang, "IRNet: INVANETs performance prediction via spatio-temporal graph attention networks," in *Proc. IEEE 21st Int. Conf. Mobile Ad-Hoc Smart Syst. (MASS)*, Sep. 2024, pp. 109–117.

[267] S. Huang, H. Zhang, X. Wang, M. Chen, J. Li, and V. C. M. Leung, "Fine-grained spatio-temporal distribution prediction of mobile content delivery in 5G ultra-dense networks," *IEEE Trans. Mobile Comput.*, pp. 1–14, Jan. 2022.

[268] M. Farreras, J. Paillissé, L. Fàbrega, and P. Vilà, "GNNetSlice: A GNN-based performance model to support network slicing in B5G networks," *Comput. Commun.*, vol. 232, Feb. 2025, Art. no. 108044.

[269] Y. Liu, C. She, Y. Zhong, W. Hardjawana, F.-C. Zheng, and B. Vucetic, "Graph neural networks for quality of service improvement in interference-limited ultra-reliable and low-latency communications," *IEEE Trans. Veh. Technol.*, vol. 73, no. 3, pp. 3718–3732, Mar. 2024.

[270] Y. Jiao, P.-H. Ho, X. Lu, J. Tapolcai, and L. Peng, "On board-level failure localization in optical transport networks using graph neural network," in *Proc. 20th Int. Conf. Design Reliable Commun. Netw. (DRCN)*, May 2024, pp. 54–61.

[271] M. Ye, J. Zhang, Z. Guo, and H. J. Chao, "Path-based graph neural network for robust and resilient routing in distributed traffic engineering," *IEEE J. Sel. Areas Commun.*, vol. 43, no. 2, pp. 422–436, Feb. 2025.

[272] M. Farhan, N. Shah, L. Wang, G.-M. Muntean, and H. H. Song, "RDG-TE: Link reliability-aware DRL-GNN-based traffic engineering in SDN," *Expert Syst. Appl.*, vol. 265, Mar. 2025, Art. no. 125963.

[273] G. R. S. Kumar and G. Arul Geetha, "Graph neural networks based queuing model for optimal load balancing in mobile ad hoc network," *Int. J. Commun. Syst.*, vol. 37, no. 17, p. 5922, Nov. 2024.

[274] S. Messinis, O. E. Marai, N. E. Protonotarios, T. Taleb, and N. Doulamis, "Roads digital twin: Predictive situational awareness using 360° video streaming and graph neural networks," *IEEE Trans. Veh. Technol.*, vol. 74, no. 4, pp. 5558–5571, Apr. 2025.

[275] Z. Xiaoyu, Q. Peihan, L. Qi, D. Yuanlei, Z. Shilian, and L. Zan, "A graph-based semi-supervised approach for few-shot class-incremental modulation classification," *China Commun.*, vol. 21, no. 11, pp. 88–103, Nov. 2024.

[276] J. Zeng, S. Lin, Y. Ai, G. Wan, W. Luo, and Q. Chen, "Green communication for OverGNN enabled heterogeneous ultra-dense networks," *IET Commun.*, vol. 19, no. 1, p. 70005, Jan. 2025.

[277] D. Yan, Y. Dai, L. Lyu, and N. Cheng, "Edge-featured graph attention network for power control in cell-free and D2D heterogeneous networks," in *Proc. 16th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Oct. 2024, pp. 678–683.

[278] Z. Huang, Z. Wang, and S. Chen, "Sub-6 GHz assisted mmWave hybrid beamforming with heterogeneous graph neural network," *IEEE Trans. Commun.*, vol. 72, no. 11, pp. 6917–6928, Nov. 2024.

[279] S. Shrestha, X. Fu, and M. Hong, "Optimal solutions for joint beamforming and antenna selection: From branch and bound to graph neural imitation learning," *IEEE Trans. Signal Process.*, vol. 71, pp. 831–846, 2023.

[280] B.-J. Chen, R. Y. Chang, F.-T. Chien, and H. V. Poor, "Graph neural network-based joint beamforming for hybrid relay and reconfigurable intelligent surface aided multiuser systems," *IEEE Wireless Commun. Lett.*, vol. 12, no. 10, pp. 1811–1815, Oct. 2023.

[281] Y. Li and Y.-F. Liu, "HPE transformer: Learning to optimize multi-group multicast beamforming under nonconvex QoS constraints," *IEEE Trans. Commun.*, vol. 72, no. 9, pp. 5581–5594, Sep. 2024.

[282] W. Deng, Y. Liu, M. Li, and M. Lei, "GNN-aided user association and beam selection for mmWave-integrated heterogeneous networks," *IEEE Wireless Commun. Lett.*, vol. 12, no. 11, pp. 1836–1840, Jul. 2023.

[283] Z. Song, Y. Lu, X. Chen, B. Ai, Z. Zhong, and D. Niyato, "A deep learning framework for physical-layer secure beamforming," *IEEE Trans. Veh. Technol.*, vol. 73, no. 12, pp. 19844–19849, Dec. 2024.

[284] H. Ji, X. Wu, Z. Zeng, and C. Chen, "Learning load balancing with GNN in MPTCP-enabled heterogeneous networks," 2024, *arXiv:2410.17118*.

[285] D. Chen, W. Zhang, D. Gao, D. Yang, and H. Zhang, "GFlow: GNN-based optimal flow scheduling for multipath transmission with link overlapping," *IEEE Trans. Netw. Sci. Eng.*, vol. 11, no. 6, pp. 6244–6258, Nov. 2024.

[286] B. Li, W. Liu, W. Xie, N. Zhang, and Y. Zhang, "Adaptive digital twin for UAV-assisted integrated sensing, communication, and computation networks," *IEEE Trans. Green Commun. Netw.*, vol. 7, no. 4, pp. 1996–2009, Dec. 2023.

[287] L. U. Khan, Z. Han, W. Saad, E. Hossain, M. Guizani, and C. S. Hong, "Digital twin of wireless systems: Overview, taxonomy, challenges, and opportunities," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 4, pp. 2230–2254, 2022.

[288] A. Isah, H. Shin, I. Aliyu, R. M. Sulaiman, and J. Kim, "Graph neural network for digital twin network: A conceptual framework," in *Proc. Int. Conf. Artif. Intell. Inf. Commun. (ICAIIC)*, Feb. 2024, pp. 1–5.

[289] P. Almasan et al., "Digital twin network: Opportunities and challenges," 2022, *arXiv:2201.01144*.

[290] Y. Zhu, G. Liu, N. Li, M. Kang, Q. Zheng, and T. Chen, "DT-GCN based digital twin for 6G self-intelligent network," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC Workshops)*, China, Aug. 2024, pp. 299–304.

[291] M. Ferriol-Galmés et al., "Building a digital twin for network optimization using graph neural networks," *Comput. Netw.*, vol. 217, Nov. 2022, Art. no. 109329.

[292] P. Yu et al., "Digital twin driven service self-healing with graph neural networks in 6G edge networks," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 11, pp. 3607–3623, Nov. 2023.

[293] Y. Zhang, J. Hu, and G. Min, "Digital twin-driven intelligent task offloading for collaborative mobile edge computing," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 10, pp. 3034–3045, Oct. 2023.

[294] L. Tang, Y. Du, Q. Liu, J. Li, S. Li, and Q. Chen, "Digital-twin-assisted resource allocation for network slicing in industry 4.0 and beyond using distributed deep reinforcement learning," *IEEE Internet Things J.*, vol. 10, no. 19, pp. 16989–17006, Oct. 2023.

[295] Y. Lu, S. Maharjan, and Y. Zhang, "Adaptive edge association for wireless digital twin networks in 6G," *IEEE Internet Things J.*, vol. 8, no. 22, pp. 16219–16230, Nov. 2021.

[296] Y. Ma, R. Du, A. Abdelraouf, K. Han, R. Gupta, and Z. Wang, "Driver digital twin for online recognition of distracted driving behaviors," *IEEE Trans. Intell. Vehicles*, vol. 9, no. 2, pp. 3168–3180, Feb. 2024.

[297] B. Wang, Z. Li, Z. Xu, Z. Sun, and K. Tian, "Digital twin modeling for structural strength monitoring via transfer learning-based multi-source data fusion," *Mech. Syst. Signal Process.*, vol. 200, Oct. 2023, Art. no. 110625.

[298] W. Jiang et al., "Graph neural networks for routing optimization: Challenges and opportunities," *Sustainability*, vol. 16, no. 21, p. 9239, Oct. 2024.

[299] R. Zhang, Y. Lu, W. Chen, B. Ai, and Z. Ding, "Model-based GNN enabled energy-efficient beamforming for ultra-dense wireless networks," *IEEE Trans. Wireless Commun.*, vol. 24, no. 4, pp. 3333–3345, Apr. 2025.

[300] J. Xu, Y. Wang, B. Zhang, and J. Ma, "A graph reinforcement learning based SDN routing path selection for optimizing long-term revenue," *Future Gener. Comput. Syst.*, vol. 150, pp. 412–423, Jan. 2024.

[301] H. Ma, Y. Rong, and J. Huang, "Graph neural networks: Scalability," in *Graph Neural Networks: Foundations, Frontiers, and Applications*. Singapore: Springer, Jan. 2022.

[302] Z. Zhao, G. Verma, A. Swami, and S. Segarra, "Distributed link sparsification for scalable scheduling using graph neural networks," *IEEE Trans. Wireless Commun.*, vol. 25, pp. 3879–3893, 2026.

[303] H. T. Kose, J. Nunez-Yanez, R. Piechocki, and J. Pope, "A survey of computationally efficient graph neural networks for reconfigurable systems," *Information*, vol. 15, no. 7, p. 377, Jun. 2024.

[304] H. Tang and Y. Liu, "Towards understanding generalization of graph neural networks," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2023, pp. 33674–33719.

[305] H. Li, X. Wang, Z. Zhang, and W. Zhu, "Out-of-distribution generalization on graphs: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 47, no. 11, pp. 10490–10512, Nov. 2025.

[306] S. Liu and K. Ding, "Beyond generalization: A survey of out-of-distribution adaptation on graphs," 2024, *arXiv:2402.11153*.

[307] F. Zhou, S. Liu, X. Liu, R. Yu, and G. Xue, "Traffic engineering in large-scale networks with generalizable graph neural networks," *IEEE Trans. Netw.*, vol. 34, pp. 3137–3149, 2026.

[308] L. Wang et al., "FlexGraph: A flexible and efficient distributed framework for GNN training," in *Proc. 16th Eur. Conf. Comput. Syst.*, Apr. 2021, pp. 67–82.

[309] M. Farreras, P. Soto, M. Camelo, L. Fàbrega, and P. Vilà, "Improving network delay predictions using GNNs," *J. Netw. Syst. Manage.*, vol. 31, no. 4, p. 65, Oct. 2023.

[310] H. Zhou, W. Xia, G. Zheng, and H. Zhu, "Graph neural network-based continual learning for resource allocation in dynamic wireless environments," *IEEE Trans. Veh. Technol.*, vol. 74, no. 12, pp. 19125–19140, Dec. 2025.

[311] J. Gao, J. Hu, G. Min, and F. Hao, "XDGNN: Efficient distributed GNN training via explanation-guided subgraph expansion," *IEEE Trans. Parallel Distrib. Syst.*, vol. 36, no. 11, pp. 2354–2365, Nov. 2025.

[312] S. Wang, M. A. Qureshi, L. Miralles-Pechuán, T. Huynh-The, T. R. Gadekallu, and M. Liyanage, "Explainable AI for 6G use cases: Technical aspects and research challenges," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 2490–2540, 2024.

[313] J. Guo and C. Yang, "When attention is beneficial for learning wireless resource allocation efficiently?," 2025, *arXiv:2507.02427*.

[314] F. Rezazadeh et al., "Toward explainable reasoning in 6G: A proof of concept study on radio resource allocation," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 6239–6260, 2024.

[315] Y. Li, J. Zhou, B. Zheng, N. Shafiabady, and F. Chen, "Reliable and faithful generative explainers for graph neural networks," *Mach. Learn. Knowl. Extraction*, vol. 6, no. 4, pp. 2913–2929, Dec. 2024.

[316] S. Akkas, A. Devarakonda, and A. Azad, "DistShap: Scalable GNN explanations with distributed Shapley values," 2025, *arXiv:2506.22668*.

[317] H. Zhao, K. Liu, M. Liu, S. Garg, and M. Alrashoud, "Intelligent beamforming for UAV-assisted IIoT based on hypergraph inspired explainable deep learning," *IEEE Trans. Consum. Electron.*, vol. 70, no. 1, pp. 1972–1982, Feb. 2024.

[318] Z. Feng, R. Wang, T. Wang, M. Song, S. Wu, and S. He, "A comprehensive survey of dynamic graph neural networks: Models, frameworks, benchmarks, experiments and challenges," *IEEE Trans. Knowl. Data Eng.*, vol. 38, no. 1, pp. 26–46, Jan. 2026.

[319] Y. Lu, Y. Li, R. Zhang, W. Chen, B. Ai, and D. Niyato, "Graph neural networks for wireless networks: Graph representation, architecture and evaluation," *IEEE Wireless Commun.*, vol. 32, no. 1, pp. 150–156, Feb. 2025.

[320] M. Marwani and G. Kaddoum, "Graph neural networks approach for joint wireless power control and spectrum allocation," *IEEE Trans. Mach. Learn. Commun. Netw.*, vol. 2, pp. 717–732, 2024.

[321] T. Dash, A. Srinivasan, and L. Vig, "Incorporating symbolic domain knowledge into graph neural networks," *Mach. Learn.*, vol. 110, no. 7, pp. 1609–1636, Jul. 2021.

[322] R. Alansary and N. Ehab, "Hybrid neuro-symbolic learning agents: A graph-theoretic approach to interpretable reinforcement learning," in *Proc. Int. Conf. Practical Appl. Agents Multi-Agent Syst. (PAAMS)*. Cham, Switzerland: Springer, 2025, pp. 107–118.

[323] L. Li, "Research on digital twin technology for future 6G," *Intell. Converged Netw.*, vol. 6, no. 1, pp. 20–40, Mar. 2025.

[324] S. Norouzi, M. Rahmani, Y. Chu, T. Braun, K. Chowdhury, and A. Burr, "Lightweight graph neural networks for enhanced 5G NR channel estimation," in *Proc. IEEE 36th Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2025, pp. 1–7.

[325] S. Norouzi, E. Samikwa, M. Rahmani, T. Braun, and A. Burr, "Decentralized federated learning for GNN-based channel estimation with DM-RS in O-RAN," in *Proc. IEEE Int. Conf. Mach. Learn. Commun. Netw. (ICMLCN)*, May 2025, pp. 1–7.

[326] M. Soltani, V. Pourahmadi, A. Mirzaei, and H. Sheikhzadeh, "Deep learning-based channel estimation," *IEEE Commun. Lett.*, vol. 23, no. 4, pp. 652–655, Apr. 2019.

[327] K. Tan, D. Bremner, J. L. Kernec, Y. Sambo, L. Zhang, and M. A. Imran, "Graph neural network-based cell switching for energy optimization in ultra-dense heterogeneous networks," *Sci. Rep.*, vol. 12, no. 1, p. 21581, Dec. 2022.

[328] H. Ahmadi et al., "Toward sustainability in 6G and beyond: Challenges and opportunities of open RAN," *IEEE Commun. Standards Mag.*, vol. 9, no. 3, pp. 126–135, Sep. 2025.

[329] Y. Feng and Q. Qian, "PPFedGNN: An efficient privacy-preserving federated graph neural network method for social network analysis," *IEEE Trans. Computat. Social Syst.*, vol. 12, no. 6, pp. 4481–4500, Dec. 2025.

[330] X. Tang et al., "Federated graph neural network for privacy-preserved supply chain data sharing," *Appl. Soft Comput.*, vol. 168, Jan. 2025, Art. no. 112475.

[331] B. Hua and H. Xi, "A privacy preserving intrusion detection framework for IIoT in 6G networks using homomorphic encryption and graph neural networks," *Sci. Rep.*, vol. 16, no. 1, Dec. 2025.

[332] S. Wang, J. Liu, H. Xu, C. Tang, Q. Ma, and L. Huang, "Towards communication-efficient decentralized federated graph learning over non-IID data," *IEEE Trans. Mobile Comput.*, vol. 25, no. 5, pp. 6929–6947, May 2026.

[333] S. Yang, G. Ding, Z. Chen, and J. Si Yang, "GART: Graph neural network-based adaptive and robust task scheduler for heterogeneous distributed computing," *IEEE Access*, vol. 13, pp. 200196–200216, 2025.

[334] H. Yang, M. Wang, L. Dai, Y. Wu, and J. Du, "Federated graph neural networks for heterogeneous graphs with data privacy and structural consistency," in *Proc. 5th Int. Conf. Comput. Sci. Blockchain (CCSB)*, Aug. 2025, pp. 311–315.

[335] M. Shao, R. Zhang, and L. Yang, "Graph neural network-based task offloading and resource allocation for scalable vehicular networks," *IET Commun.*, vol. 19, no. 1, p. 70064, Jan. 2025.

[336] Z. Zhan, H. Zhou, and H. Haddadi, "Poster: Enhancing GNN robustness for network intrusion detection via agent-based analysis," 2025, *arXiv:2506.20806*.

[337] Y. Luo, J. Chen, S. Qin, and D. Wu, "Dual graph learning for multivariate time series anomaly detection in IoUT," *IEEE Trans. Netw. Sci. Eng.*, vol. 13, pp. 1632–1646, 2026.